

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 18:57:14 ; Search time 33.75 Seconds
(without alignments)
145.793 Million cell updates/sec

Title: US-09-868-885b-3

Perfect score: 172
Sequence: 1 SRAHQHSEMTDPINPAWAGRGIRPVGRF 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq 19Jun03.*

1: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	172	100.0	31	AAW31371	Bovine G protein-c
2	172	100.0	31	AAW97218	Bovine pituitary-d
3	172	100.0	31	AAW87613	Bovine 19P2 ligand
4	172	100.0	31	AAW95188	Bovine pituitary-d
5	172	100.0	31	AAW951347	Bovine oxytocin se
6	172	100.0	31	AAW49290	19P2 ligand peptid
7	172	100.0	31	AAW49298	19P2 ligand peptid
8	172	100.0	31	AAG62516	Bovine CRH releasi
9	172	100.0	31	AAE26399	Bovine PrRP-31 pep

10	172	100.0	31	24	ABU60825	Peptide production
11	172	100.0	31	24	ABU60831	Peptide production
12	172	100.0	32	18	AAW31372	Bovine G protein-c
13	172	100.0	32	20	AAW95189	Bovine pituitary-d
14	172	100.0	32	21	AAW10348	Bovine oxytocin se
15	172	100.0	32	22	AAG62517	Bovine CRH releasi
16	172	100.0	32	24	ABU60832	Peptide production
17	172	100.0	33	18	AAW31373	Bovine G protein-c
18	172	100.0	33	20	AAW95190	Bovine pituitary-d
19	172	100.0	33	21	AAW10349	Bovine oxytocin se
20	172	100.0	33	21	AAW49297	19P2 ligand peptid
21	172	100.0	33	22	AAG62518	Bovine CRH releasi
22	172	100.0	33	24	ABU60833	Peptide production
23	172	100.0	98	18	AAW31382	Bovine genome deri
24	172	100.0	98	18	AAW31368	Bovine G protein-c
25	172	100.0	98	20	AAW97224	Bovine genome-driv
26	172	100.0	98	20	AAW97217	Bovine pituitary-d
27	172	100.0	98	20	AAW95187	Bovine genome-deri
28	172	100.0	98	21	AAW10346	Bovine oxytocin se
29	172	100.0	98	21	AAW10353	Bovine oxytocin se
30	172	100.0	98	22	AAG62515	Bovine CRH releasi
31	172	100.0	98	22	AAG62522	Bovine CRH releasi
32	166	96.5	31	22	AAW90995	Prolactin releasin
33	164	95.3	31	22	AAW73370	bPrp31 peptide, u
34	163	94.8	31	18	AAW31384	Rat type G protein
35	163	94.8	31	20	AAW97233	Rat type ligand po
36	163	94.8	31	20	AAW87614	Rat 19P2 ligand.
37	163	94.8	31	20	AAW95173	Murine pituitary-d
38	163	94.8	31	20	AAW95174	Murine pituitary-d
39	163	94.8	31	21	AAW10355	Rat oxytocin secre
40	163	94.8	31	21	AAW87504	Rat prolactin-rele
41	163	94.8	31	21	AAW49292	19P2 ligand peptid
42	163	94.8	31	22	AAG62524	Rat CRH releasing
43	163	94.8	31	22	AAW90993	Prolactin releasin
44	163	94.8	31	23	AAE26400	Rat PrRP-31 peptid
45	163	94.8	31	24	ABU60826	Peptide production

ALIGNMENTS

RESULT 1

AAW31371
ID AAW31371 standard; Peptide; 31 AA.

XX AAW31371;

AC AAW31371;

XX 06-APR-1998 (first entry)

DT Bovine G protein-coupled receptor ligand peptide fragment 1.

XX G protein-coupled receptor; ligand binding; pharmaceutical;
modulator; pituitary; central nervous system; pancreas; prophylactic;
therapeutic agent.

OS Bos taurus.

XX MO9724436-A2.

PD 10-JUL-1997.

XX 26-DEC-1996; 96WO-JP03821.

XX 18-SEP-1996; 96JP-0246573.

PR 28-DEC-1995; 95JP-0343371.

PR 15-MAR-1996; 96JP-0059419.

PR 12-AUG-1996; 96JP-0211805.

XX (TAKE) TAKEDA CHEM IND LTD.

XX Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;

PI Kawamata Y, Kitada C;

XX

DR WPI; 1997-363672/33.
 DR N-PSDB; AAV02394.
 PT Ligand peptide for G protein-coupled receptor - acts by modulating
 PT function in the central nervous system, pancreas and pituitary gland
 XX
 PS
 XX Claim 2; Page 160; 258pp; English.
 XX
 CC This sequence represents a peptide fragment of a novel bovine pituitary
 CC derived ligand corresponding to amino acid residues 23 to 53 of the
 CC sequence in AAW31368 and is used in an assay to monitor ligand binding
 CC to the G protein-coupled receptor protein. Pharmaceutical compositions
 CC containing this ligand may be used as a pituitary function modulator, a
 CC central nervous system modulator or a pancreatic function modulator.
 CC This ligand could have specific applications as a prophylactic or
 CC therapeutic agent for dementia, depression, hyperkinetic syndrome,
 CC disturbance of consciousness, anxiety syndrome, schizophrenia, trauma,
 CC growth hormone secretory disease, hyper- and polyphagia,
 CC hypercholesterolaemia, hyperglycaemia, hyperlipidaemia,
 CC hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease,
 CC Turner's syndrome, neurosis, rheumatoid arthritis, spinal injury,
 CC transient brain ischaemia, amyotrophic lateral sclerosis, acute
 CC myocardial infarction, spinocerebellar degeneration, bone fracture,
 CC trauma, atopic dermatitis, osteoporosis, asthma, epilepsy, infertility
 CC and/or oligogalactia. Assays can also be developed to screen compounds
 CC which are capable of altering the binding activity of the ligand thus
 CC affecting activation of the G protein-coupled receptor protein.
 XX
 SX Sequence 31 AA;
 Query Match 100.0%; Score 172; DB 18; Length 31;
 Best Local Similarity 100.0%; Pred. No. 1.1e-18;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SRAHQHSMETPTDINPAWYAGRGIRPVGRF 31
 |||||
 DB 1 SRAHQHSMETPTDINPAWYAGRGIRPVGRF 31
 |||||
 RESULT 2
 AAW97218
 ID AAW97218 standard; peptide; 31 AA.
 AC AAW97218;
 XX
 DT 06-MAY-1999 (first entry)
 XX
 DE Bovine pituitary-derived ligand polypeptide fragment.
 XX
 KW Bovine pituitary-derived ligand; modulation; prolactin secretion;
 KW G protein-coupled receptor; GPCR; hypocoovarianism; gonocyst cacogenesis;
 KW menopausal syndrome; euthyroid; hypometabolism; lactation;
 KW pituitary adenomatosis; brain tumour; emmenopathy; autoimmune disease;
 KW prolactinoma; infertility; impotence; amenorrhea; galactorrhea;
 KW acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;
 KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;
 KW contraceptive; placental function; choriocarcinoma; hydatid mole;
 KW irruption mole; abortion; unthrifty fetus; abnormal saccharometabolism;
 KW abnormal lipidmetabolism; oxytocia.
 XX
 OS Bos sp.
 XX
 PN WO9858962-A1.
 XX
 PD 30-DEC-1998.
 XX
 PF 22-JUN-1998; 98WO-JP02765.
 XX
 PR 23-JUN-1997; 97JP-0165437.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX
 PI Fujii R, Hinuma S, Kawanata Y, Matsumoto H;
 WPI; 1999-105614/09.
 Use of G protein-coupled receptor ligands - for modulating prolactin
 secretion or placental function, e.g. for treating menopausal
 syndrome, tumours, autoimmune disease or abnormal pregnancy
 Claim 3; Page 135; 241pp; English.
 The present sequence represents a bovine pituitary-derived ligand
 fragment. It is used in the course of the invention. The specification
 describes an agent for modulating prolactin secretion which comprises a
 ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)
 protein. The agents for promoting prolactin secretion can be used for
 treating or preventing hypocoovarianism, gonocyst cacogenesis, menopausal
 syndrome, euthyroid or hypometabolism. They can be used for promoting
 lactation in a domestic mammal and as an aphrodisiac. The agents for
 inhibiting prolactin secretion can be used for treating or preventing
 pituitary adenomatosis, brain tumour, emmenopathy, autoimmune disease,
 prolactinoma, infertility, impotence, amenorrhea, galactorrhea,
 acromegaly, Chiari-Frommel syndrome, Argonz-del Castillo syndrome,
 Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dyszoospermia.
 The inhibitory agents can also be used as contraceptives. The agents for
 modulating placental function can be used for treating or preventing
 choriocarcinoma, hydatid mole, irruption mole, abortion, unthrifty fetus,
 abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.

XX CC This is the amino acid sequence of the bovine pituitary G
 CC protein-coupled receptor ligand 19p2L. A method suitable for
 CC commercial high-level production of 19p2L comprises expressing
 CC the ligand in host cells as a recombinant fusion protein e.g. with
 CC human basic fibroblast growth factor (see AAV83792-93) that has
 CC been modified to include an N-terminal cysteine residue. The
 CC ligand is released from the fusion by cyanylation followed by
 CC ammonolysis. 19p2L has prolactin secretion-stimulating and (at
 CC high doses) prolactin secretion-inhibiting properties. It can be
 CC used in the treatment and prevention of various diseases including:
 CC senile dementia, cerebrovascular dementia, and dementia associated
 CC with: neurological disorders (e.g. Alzheimer's disease, Parkinson's
 CC disease, Pick's disease, Huntington's disease), infectious diseases
 CC (e.g. Creutzfeldt-Jakob's), endocrine or metabolic disease or
 CC toxicosis (e.g. hypothyroidism, vitamin B12 deficiency, alcoholism,
 CC intoxication by drugs, metal and organic compounds), tumorigenic
 CC diseases (e.g. brain tumour), traumatic diseases (e.g. chronic
 CC subarachnoidal haemorrhage, and other types of dementia, depression,
 CC hyperactive child syndrome (microencephalopathy) and disturbance of
 CC consciousness. It is also useful for prevention and treatment of
 CC diseases associated with prolactin hypo and hypersecretion
 CC respectively, including: hyperprolactinaemia, pituitary adenoma,
 CC breast cancer, infertility, impotence and autoimmune disease
 CC (hypersecretion disorders), and seminal vesicle hypoplasia,
 CC osteoporosis, menopausal syndrome and renal failure (hyposecretion
 CC disorders). The 19p2 polypeptide/amide is also useful as a test
 CC reagent for study of the prolactin secretory function or as a
 CC lactagogue in mammalian farm animals.

XX CC Sequence 31 AA;

Query Match 100.0%; Score 172; DB 20; Length 31;
 Best Local Similarity 100.0%; Pred. No. 1.1e-18;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSMEIRTPDINPAWYAGRGIRPVGRF 31
 |||||
 DB 1 SRAHQHSMEIRTPDINPAWYAGRGIRPVGRF 31
 |||||

RESULT 4
 AAW95188
 ID AAW95188 standard; peptide; 31 AA.
 AC AAW95188;
 XX
 DT 10-MAR-1999 (first entry)
 XX
 DE Bovine pituitary-derived ligand polypeptide fragment.
 XX
 KW Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;
 KW GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas;
 KW tissue; screen; therapeutic; binding; senile dementia; ligand; murine;
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;
 KW Creutzfeld-Jakob disease; poisoning; schizophrenia; growth hormone;
 KW secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;
 KW gene therapy; transgenic animal; bovine.

OS Bos sp.
 XX
 PN WO9849295-A1.
 XX
 PD 05-NOV-1998.
 XX
 PF 27-APR-1998; 98WO-JP01923.
 XX
 PR 28-APR-1997; 97JP-0109974.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX
 PI Fukusumi S, Hinuma S;
 XX
 DR WPI; 2000-452298/39.
 XX

DR WPI; 1999-009423/01.
 XX
 PT New polypeptide ligand for orphan G protein coupled receptors - used
 PT for treating disorders of central nervous system, pituitary and
 PT pancreas, and for drug screening
 XX
 PS Example 19; Page 150; 206pp; English.
 XX
 CC The invention relates to a murine pituitary-derived ligand polypeptide
 CC which is a ligand for the G-protein coupled orphan receptor designated
 CC GPR10 (human) or UHR-1 (rat). Cells transformed with a vector containing
 CC the ligand polypeptide encoding DNA are used to produce a recombinant
 CC ligand polypeptide. The ligand polypeptide, and its fragments, modulate
 CC function of the pituitary, central nervous system, pancreas and other
 CC tissues and can be used to screen for agents that modulate binding of the
 CC polypeptide to the receptor; to quantify the amount of receptor in a
 CC sample and to raise antibodies. They may also be used therapeutically,
 CC e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's
 CC diseases; Creutzfeld-Jakob disease; poisoning by heavy metals or drugs;
 CC diabetes; schizophrenia; disorders of growth hormone secretion; cancer;
 CC rheumatoid arthritis; epilepsy and many others, also to improve post-
 CC operative nutritional status and as vasopressor. Transgenic animals
 CC carrying the ligand polypeptide encoding DNA or its mutin are used to
 CC study the function of the polypeptide-expressing genes, as models of
 CC disease, for drug screening and as source of cell lines. The ligand
 CC polypeptide DNA is used as a source of probes and primers; to identify
 CC related sequences; in receptor-binding assays; for production of Ab and
 CC antisera; in drug development; for gene therapy and to develop transgenic
 CC animals. The present sequence represents a bovine genome-derived ligand
 CC polypeptide fragment which is similar to the murine ligand-polypeptide.

XX CC Sequence 31 AA;

Query Match 100.0%; Score 172; DB 20; Length 31;
 Best Local Similarity 100.0%; Pred. No. 1.1e-18;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSMEIRTPDINPAWYAGRGIRPVGRF 31
 |||||
 DB 1 SRAHQHSMEIRTPDINPAWYAGRGIRPVGRF 31
 |||||

RESULT 5
 AAB10347
 ID AAB10347 standard; peptide; 31 AA.
 XX
 AC AAB10347;
 XX
 DT 24-NOV-2000 (first entry)
 XX
 DE Bovine oxytocin secretion promoting peptide SEQ ID NO: 3.
 XX
 KW Bovine; oxytocin secretion promoter; G protein-coupled receptor protein;
 KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;
 KW caesarean section; artificial fertilization; galactostasis; goat; pig;
 KW veterinary medicine; milk production.

OS Bos taurus.
 XX
 PN WO200038704-A1.
 XX
 PD 06-JUL-2000.
 XX
 PF 22-DEC-1999; 99WO-JP07199.
 XX
 PR 25-DEC-1998; 98JP-0369585.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX
 PI Matsumoto H, Kitada C, Hinuma S;
 XX
 DR WPI; 2000-452298/39.
 XX

PT Physiologically-active polypeptide recognized as ligand by G
 PT protein-coupled receptor protein, for promoting secretion of oxytocin,
 PT as drugs for diseases relating to oxytocin secretion and in veterinary
 PT medicine -
 XX
 XX
 PS Claim 3; Page 50; 72pp; Japanese.
 CC
 CC This invention describes a novel oxytocin secretion-regulating agent
 CC which contains a ligand peptide or its salt for the G protein-coupled
 CC receptor protein. It is useful in the form of drugs for ameliorating,
 CC preventing and treating diseases relating to oxytocin secretion e.g.
 CC weak pains and atonic bleeding, before and after expulsion of placenta,
 CC uterine recovery failure, caesarean section, stoppage of artificial
 CC fertilization or galactostasis and is also applicable in veterinary
 CC medicine for promoting milk production in cow, goat and pig. This
 CC sequence represents a bovine peptide which acts as an oxytocin secretion
 CC promoter.
 XX
 XX
 SQ Sequence 31 AA;
 Query Match 100.0%; Score 172; DB 21; Length 31;
 Best Local Similarity 100.0%; Pred. No. 1.1e-18;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 SRAHQHSMEIRTPDINPAWYAGRGIRPVGRF 31
 Db 1 SRAHQHSMEIRTPDINPAWYAGRGIRPVGRF 31
 RESULT 6
 AAY49290
 ID AAY49290 standard; peptide; 31 AA.
 XX
 AC AAY49290;
 XX
 DT 22-FEB-2000 (first entry)
 XX
 DE 19P2 ligand peptide fragment.
 XX
 KW Monoclonal antibody; 19P2 ligand; diagnosis; prolactin secretion;
 KW pituitary; regulatory mechanism; central nervous system; pancreatic.
 OS Bos sp.
 XX
 XX Key Location/Qualifiers
 FH Modified-site 31
 FT /note= "C-terminal amide"
 FT
 XX WO9960112-A1.
 XX
 PD 25-NOV-1999.
 XX
 XX 20-MAY-1999; 99WO-JP02650.
 XX
 PR 21-MAY-1998; 98JP-0140293.
 XX
 XX (TAKE) TAKEDA CHEM IND LTD.
 XX
 PI Matsumoto H, Kitada C, Hinuma S;
 XX
 XX WPI; 2000-039381/03.
 XX
 XX New monoclonal antibodies, useful in diagnosis, as drugs and in
 XX studying diseases related to ligand abnormality -
 XX
 PS Disclosure; Page 26; 73pp; Japanese.
 CC
 CC The invention provides a monoclonal antibody which has a specific
 CC reaction with the part peptide of the C-terminal of 19P2 ligand or its
 CC derivative. The antibodies can be used in diagnosis or to treat or
 CC prevent diseases associated with abnormality in the pituitary function
 CC regulatory mechanism (e.g. promotion of prolactin secretion), central
 CC nervous regulatory mechanism, and pancreatic function regulatory

CC mechanism. The antibody-based immunoassay can also be applied in
 CC clarifying the physiological functions of the ligand and its derivative.
 CC Sequences AAY49290-302 represent peptide fragments of the 19P2 ligand.
 XX
 XX
 SQ Sequence 31 AA;
 Query Match 100.0%; Score 172; DB 21; Length 31;
 Best Local Similarity 100.0%; Pred. No. 1.1e-18;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 SRAHQHSMEIRTPDINPAWYAGRGIRPVGRF 31
 Db 1 SRAHQHSMEIRTPDINPAWYAGRGIRPVGRF 31
 RESULT 7
 AAY49298
 ID AAY49298 standard; peptide; 31 AA.
 XX
 AC AAY49298;
 XX
 DT 22-FEB-2000 (first entry)
 XX
 DE 19P2 ligand peptide fragment.
 XX
 KW Monoclonal antibody; 19P2 ligand; diagnosis; prolactin secretion;
 KW pituitary; regulatory mechanism; central nervous system; pancreatic.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH Modified-site 31
 FT /note= "C-terminal amide"
 FT
 XX WO9960112-A1.
 XX
 PD 25-NOV-1999.
 XX
 XX 20-MAY-1999; 99WO-JP02650.
 XX
 PR 21-MAY-1998; 98JP-0140293.
 XX
 XX (TAKE) TAKEDA CHEM IND LTD.
 XX
 PI Matsumoto H, Kitada C, Hinuma S;
 XX
 XX WPI; 2000-039381/03.
 XX
 XX New monoclonal antibodies, useful in diagnosis, as drugs and in
 XX studying diseases related to ligand abnormality -
 XX
 PS Disclosure; Page 27; 73pp; Japanese.
 CC
 CC The invention provides a monoclonal antibody which has a specific
 CC reaction with the part peptide of the C-terminal of 19P2 ligand or its
 CC derivative. The antibodies can be used in diagnosis or to treat or
 CC prevent diseases associated with abnormality in the pituitary function
 CC regulatory mechanism (e.g. promotion of prolactin secretion), central
 CC nervous regulatory mechanism, and pancreatic function regulatory
 CC mechanism. The antibody-based immunoassay can also be applied in
 CC clarifying the physiological functions of the ligand and its derivative.
 CC Sequences AAY49290-302 represent peptide fragments of the 19P2 ligand.
 XX
 XX
 SQ Sequence 31 AA;
 Query Match 100.0%; Score 172; DB 21; Length 31;
 Best Local Similarity 100.0%; Pred. No. 1.1e-18;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 SRAHQHSMEIRTPDINPAWYAGRGIRPVGRF 31
 Db 1 SRAHQHSMEIRTPDINPAWYAGRGIRPVGRF 31

RESULT 8
AAG62516
ID AAG62516 standard; peptide; 31 AA.
XX AC AAG62516;
XX DT 24-AUG-2001 (first entry)
XX DE Bovine CRH releasing protein related peptide SEQ ID NO: 3.
XX KW Cow; corticotrophin releasing hormone; CRH; G protein receptor ligand;
XX KW analgesic; hyperaldosteronism; hypercortisolemia; hypoadrenocorticism;
XX KW Addison's disease; adrenal gland hyperfunction; obesity.
XX OS Bos taurus.
XX PN WO200135984-A1.
XX PD 25-MAY-2001.
XX PF 17-NOV-2000; 2000WO-JP08119.
XX PR 18-NOV-1999; 99JP-0327900.
XX PR 26-SEP-2000; 2000JP-0297073.
XX PA (TAKE) TAKEDA CHEM IND LTD.
XX PI Kitada C, Matsumoto H, Hinuma S;
XX DR WPI; 2001-355552/37.
XX PT Use of G protein receptor ligand or peptide for controlling
XX PT corticotrophin releasing hormone secretion -
XX PS Claim 3. Page 63; 90pp; Japanese.
XX CC The present sequence describes a method of controlling the secretion of
XX CC corticotrophin releasing hormone (CRH), involving the use of a G protein
XX CC receptor ligand. This can be used to control the secretion of CRH and is
XX CC useful as an analgesic or for treating, preventing or ameliorating
XX CC diseases associated with CRH secretion such as hyperaldosteronism,
XX CC hypercortisolemia, secondary or chronic hypoadrenocorticism, Addison's
XX CC disease (including boredom, nausea, pigmentation, hypogonadism, hair
XX CC loss, and hypotension), adrenal gland hypofunction and obesity. The
XX CC present sequence is a peptide used in the exemplification of the
XX CC invention.
XX SQ Sequence 31 AA;
Query Match 100.0%; Score 172; DB 22; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.1e-18;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SRAHQHSMETPTDINPAWYAGRGIRPVGRF 31
DB 1 SRAHQHSMETPTDINPAWYAGRGIRPVGRF 31
RESULT 9
AAE26399
ID AAE26399 standard; peptide; 31 AA.
XX AC AAE26399;
XX DT 13-DEC-2002 (first entry)
XX DE Bovine PrRP-31 peptide.
XX KW Wakefulness; sleep disorder; prolactin releasing peptide receptor;
XX KW PrRP; GPR10; therapy; epilepsy; narcolepsy; sleepiness; sleep apnoea;
XX KW insomnia; idiopathic hypersomnia; psychogenic hypersomnia; seizure;
XX KW anticonvulsant; bovine.

XX OS Bos taurus.
XX PN US2002037533-A1.
XX PD 28-MAR-2002.
XX PF 17-AUG-2001; 2001US-0932161.
XX PR 28-APR-2000; 2000US-0560915.
XX PA (CIVE/) CIVELLI O.
XX PA (LINS/) LIN S.
XX PI Civelli O, Lin S;
XX DR WPI; 2002-403931/43.
XX PT Screening for compounds useful for promoting wakefulness or sleep, and
XX PT for treating sleeping disorders, e.g. insomnia, hypersomnia or sleep
XX PT apnea, comprises administering a prolactin releasing peptide agonist or
XX PT antagonist -
XX PS Disclosure; Page 24; 35pp; English.
XX CC The present invention relates to a method of screening for compounds
XX CC for promoting wakefulness or sleep in a mammal. The method involves
XX CC administering a prolactin releasing peptide (PrRP) receptor (GPR10)
XX CC agonist or antagonist respectively and determining the ability of the
XX CC compound to promote wakefulness or sleep. The compounds identified
XX CC from the method are used in the therapy of epilepsy and other diseases
XX CC associated with absence seizures and in promoting wakefulness and sleep
XX CC in individuals having sleep disorders such as insomnia and narcolepsy.
XX CC PrRP receptor agonists may be used to treat common disorders which lead
XX CC to sleepiness, e.g. sleep apnoea, narcolepsy, idiopathic hypersomnia
XX CC and psychogenic hypersomnia. PrRP receptor antagonists are useful for
XX CC promoting sleep and for treating insomnia such as adjustment sleep
XX CC disorder and psychophysiological insomnia. The present sequence is bovine
XX CC PrRP-31 peptide.
XX SQ Sequence 31 AA;
Query Match 100.0%; Score 172; DB 23; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.1e-18;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SRAHQHSMETPTDINPAWYAGRGIRPVGRF 31
DB 1 SRAHQHSMETPTDINPAWYAGRGIRPVGRF 31
RESULT 10
ABU60825
ID ABU60825 standard; Peptide; 31 AA.
XX AC ABU60825;
XX DT 06-MAY-2003 (first entry)
XX DE Peptide production by gene recombination associated peptide #9.
XX KW Peptide production; low-molecular peptide; KiSS-1; GPR8 ligand;
XX KW gene recombination.
XX OS Bos sp.
XX PN WO200292829-A1.
XX PD 21-NOV-2002.
XX PF 16-MAY-2002; 2002WO-JP04735.
XX PR 17-MAY-2001; 2001JP-0147341.

XX PA (TAKE) TAKEDA CHEM IND LTD.
 XX PI Nishimura O, Suenaga M, Ito T, Kitada C;
 XX DR WPI; 2003-129302/12.
 XX PT Process for producing peptides e.g. Kiss-1 peptide and GPR8 ligand for
 PT subsequent applications by gene recombination technique through tandem
 PT repeats to provide precursor protein with specific cleavage sites -
 XX PS Disclosure; Page 58; 87pp; Japanese.
 XX CC The invention describes a method of producing a peptide comprising the
 CC excision of the N and C-terminals of a target peptide with enzymes or
 CC chemically through the attached cleavage sites repeated by ligation in a
 CC precursor protein. The method is for producing (low-molecular) peptides
 CC e.g. Kiss-1 peptide and GPR8 ligand for subsequent applications by the
 CC gene recombination technique through tandem repeats to provide
 CC a precursor protein with specific cleavage sites. With this method,
 CC peptide production can be carried out easily to provide large quantities
 CC of the required peptides. This is the amino acid sequence of a peptide
 CC associated with the peptide production method of the invention.
 XX SQ Sequence 31 AA;
 Query Match 100.0%; Score 172; DB 24; Length 31;
 Best Local Similarity 100.0%; Pred. No. 1.1e-18;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SRAHQHSMIEITPDINPAWYAGRGIRPVGRF 31
 DB 1 SRAHQHSMIEITPDINPAWYAGRGIRPVGRF 31
 RESULT 11
 ABU60831
 ID ABU60831 standard; Peptide; 31 AA.
 XX AC ABU60831;
 XX DT 06-MAY-2003 (first entry)
 XX DE Peptide production by gene recombination associated peptide #15.
 XX KW Peptide production; low-molecular peptide; Kiss-1; GPR8 ligand;
 XX gene recombination.
 XX OS Bos sp.
 XX PN WO200292829-A1.
 XX PD 21-NOV-2002.
 XX PF 16-MAY-2002; 2002WO-JP04735.
 XX PR 17-MAY-2001; 2001JP-0147341.
 XX PA (TAKE) TAKEDA CHEM IND LTD.
 XX PI Nishimura O, Suenaga M, Ito T, Kitada C;
 XX DR WPI; 2003-129302/12.
 XX PT Process for producing peptides e.g. Kiss-1 peptide and GPR8 ligand for
 PT subsequent applications by gene recombination technique through tandem
 PT repeats to provide precursor protein with specific cleavage sites -
 XX PS Disclosure; Page 61; 87pp; Japanese.
 XX CC The invention describes a method of producing a peptide comprising the
 CC excision of the N and C-terminals of a target peptide with enzymes or
 CC chemically through the attached cleavage sites repeated by ligation in a

CC precursor protein. The method is for producing (low-molecular) peptides
 CC e.g. Kiss-1 peptide and GPR8 ligand for subsequent applications by the
 CC gene recombination technique through tandem repeats to provide
 CC a precursor protein with specific cleavage sites. With this method,
 CC peptide production can be carried out easily to provide large quantities
 CC of the required peptides. This is the amino acid sequence of a peptide
 CC associated with the peptide production method of the invention.
 XX SQ Sequence 31 AA;
 Query Match 100.0%; Score 172; DB 24; Length 31;
 Best Local Similarity 100.0%; Pred. No. 1.1e-18;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SRAHQHSMIEITPDINPAWYAGRGIRPVGRF 31
 DB 1 SRAHQHSMIEITPDINPAWYAGRGIRPVGRF 31
 RESULT 12
 AAW31372
 ID AAW31372 standard; Peptide; 32 AA.
 XX AC AAW31372;
 XX DT 06-APR-1998 (first entry)
 XX DE Bovine G protein-coupled receptor ligand peptide fragment 2.
 XX KW G protein-coupled receptor; ligand binding; pharmaceutical;
 KW modulator; pituitary; central nervous system; pancreas; prophylactic;
 KW therapeutic agent.
 XX OS Bos taurus.
 XX PN WO9724436-A2.
 XX PD 10-JUL-1997.
 XX PF 26-DEC-1996; 96WO-JP03821.
 XX PR 18-SEP-1996; 96JP-0246573.
 XX PR 28-DEC-1995; 95JP-0343371.
 XX PR 15-MAR-1996; 96JP-0059419.
 XX PR 12-AUG-1996; 96JP-0211805.
 XX PA (TAKE) TAKEDA CHEM IND LTD.
 XX PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;
 PI Kawamata Y, Kitada C;
 XX DR WPI; 1997-363672/33.
 XX DR N-PSDB; AAV02395.
 XX PT Ligand peptide for G protein-coupled receptor - acts by modulating
 PT function in the central nervous system, pancreas and pituitary gland
 XX Claim 2; Page 160-161; 258pp; English.
 XX CC This sequence represents a peptide fragment of a novel bovine pituitary
 CC derived ligand corresponding to amino acid residues 23 to 54 of the
 CC sequence in AAW31368 and is used in an assay to monitor ligand binding
 CC to the G protein-coupled receptor protein. Pharmaceutical compositions
 CC containing this ligand may be used as a pituitary function modulator, a
 CC central nervous system modulator or a pancreatic function modulator.
 CC This ligand could have specific applications as a prophylactic or
 CC therapeutic agent for dementia, depression, hyperkinetic syndrome,
 CC disturbance of consciousness, anxiety syndrome, schizophrenia, trauma,
 CC growth hormone secretory disease, hyper- and polyphagia,
 CC hypercholesterolaemia, hyperglycaemia, hyperlipidaemia,
 CC hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease,
 CC Turner's syndrome, neurosis, rheumatoid arthritis, spinal injury,
 CC transient brain ischaemia, amyotrophic lateral sclerosis, acute

CC myocardial infarction, spinocerebellar degeneration, bone fracture,
 CC trauma, atopic dermatitis, osteoporosis, asthma, epilepsy, infertility
 CC and/or oligogalactia. Assays can also be developed to screen compounds
 CC which are capable of altering the binding activity of the ligand thus
 CC affecting activation of the G protein-coupled receptor protein.

XX Sequence 32 AA;

Query Match 100.0%; Score 172; DB 18; Length 32;
 Best Local Similarity 100.0%; Pred. No. 1.2e-18; Indels 0; Gaps 0;
 Matches 31; Conservative 0; Mismatches 0;

QY 1 SRAHQHSMEIRTPDINPAWYAGRGIRPVGRF 31
 |||||
 DB 1 SRAHQHSMEIRTPDINPAWYAGRGIRPVGRF 31
 |||||

RESULT 13

AAW95189
 ID AAW95189 standard; peptide; 32 AA.

XX AAW95189;

DT 10-MAR-1999 (first entry)

DE Bovine pituitary-derived ligand polypeptide fragment.

XX Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;
 KW GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas;
 KW tissue; screen; therapeutic; binding; senile dementia; ligand; murine;
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;
 KW Creutzfeld-Jakob disease; poisoning; schizophrenia; growth hormone;
 KW secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;
 KW gene therapy; transgenic animal; bovine.

OS Bos sp.

PN WO9849295-A1.

XX 05-NOV-1998.

PF 27-APR-1998; 98WO-JP01923.

PR 28-APR-1997; 97JP-0109974.

PA (TAKE) TAKEDA CHEM IND LTD.

PI Fukusumi S, Hinuma S;

DR WPI; 1999-009423/01.

XX New polypeptide ligand for orphan G protein coupled receptors - used
 PT for treating disorders of central nervous system, pituitary and
 PT pancreas, and for drug screening

PS Example 19; Page 150; 206pp; English.

XX The invention relates to a murine pituitary-derived ligand polypeptide
 CC which is a ligand for the G-protein coupled orphan receptor designated
 CC GPR10 (human) or UHR-1 (rat). Cells transformed with a vector containing
 CC the ligand polypeptide encoding DNA are used to produce a recombinant
 CC ligand polypeptide. The ligand polypeptide, and its fragments, modulate
 CC function of the pituitary, central nervous system, pancreas and other
 CC tissues and can be used to screen for agents that modulate binding of the
 CC polypeptide to the receptor; to quantify the amount of receptor in a
 CC sample and to raise antibodies. They may also be used therapeutically,
 CC e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's
 CC diseases; Creutzfeld-Jakob disease; poisoning by heavy metals or drugs;
 CC diabetes; schizophrenia; disorders of growth hormone secretion; cancer;
 CC rheumatoid arthritis; epilepsy and many others, also to improve post-
 CC operative nutritional status and as vasopressor. Transgenic animals
 CC carrying the ligand polypeptide encoding DNA or its mutin are used to
 CC study the function of the polypeptide-expressing genes, as models of

CC disease, for drug screening and as source of cell lines. The ligand
 CC polypeptide DNA is used as a source of probes and primers; to identify
 CC related sequences; in receptor-binding assays; for production of Ab and
 CC antisera; in drug development; for gene therapy and to develop transgenic
 CC animals. The present sequence represents a bovine genome-derived ligand
 CC polypeptide fragment which is similar to the murine ligand-polypeptide.

SQ Sequence 32 AA;

Query Match 100.0%; Score 172; DB 20; Length 32;
 Best Local Similarity 100.0%; Pred. No. 1.2e-18; Indels 0; Gaps 0;
 Matches 31; Conservative 0; Mismatches 0;

QY 1 SRAHQHSMEIRTPDINPAWYAGRGIRPVGRF 31
 |||||
 DB 1 SRAHQHSMEIRTPDINPAWYAGRGIRPVGRF 31
 |||||

RESULT 14

AAAB10348

ID AAB10348 standard; peptide; 32 AA.

XX AAB10348;

DT 24-NOV-2000 (first entry)

DE Bovine oxytocin secretion promoting peptide SEQ ID NO: 4.

XX Bovine; oxytocin secretion promoter; G protein-coupled receptor protein;
 KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;
 KW caesarean section; artificial fertilization; galactostasis; goat; pig;
 KW veterinary medicine; milk production.

OS Bos taurus.

PN WO200038704-A1.

PD 06-JUL-2000.

XX 22-DEC-1999; 99WO-JP07199.

PR 25-DEC-1998; 98JP-0369585.

XX (TAKE) TAKEDA CHEM IND LTD.

PI Matsumoto H, Kitada C, Hinuma S;

DR WPI; 2000-452298/39.

XX Physiologically-active polypeptide recognized as ligand by G
 PT protein-coupled receptor protein, for promoting secretion of oxytocin,
 PT as drugs for diseases relating to oxytocin secretion and in veterinary
 PT medicine -

PS Disclosure; Page 51; 72pp; Japanese.

XX This invention describes a novel oxytocin secretion-regulating agent
 CC which contains a ligand peptide or its salt for the G protein-coupled
 CC receptor protein. It is useful in the form of drugs for ameliorating,
 CC preventing and treating diseases relating to oxytocin secretion e.g.
 CC weak pains and atonic bleeding, before and after expulsion of placenta,
 CC uterine recovery failure, caesarean section, stoppage of artificial
 CC fertilization or galactostasis and is also applicable in veterinary
 CC medicine for promoting milk production in cow, goat and pig. This
 CC sequence represents a bovine peptide which acts as an oxytocin secretion
 CC promoter.

SQ Sequence 32 AA;

Query Match 100.0%; Score 172; DB 21; Length 32;
 Best Local Similarity 100.0%; Pred. No. 1.2e-18;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRAHQSHMEIRTPDINPAWYAGRGIRPVGRF 31
 |||||
 Db 1 SRAHQSHMEIRTPDINPAWYAGRGIRPVGRF 31

RESULT 15

AAG62517
 ID AAG62517 standard; peptide; 32 AA.
 XX
 AC AAG62517;
 XX
 DT 24-AUG-2001 (first entry)
 XX
 DE Bovine CRH releasing protein related peptide SEQ ID NO: 4.
 XX
 KW Cow; corticotrophin releasing hormone; CRH; G protein receptor ligand;
 KW analgesic; hyperaldosteronism; hypercortisolaemia; hypoadrenocorticism;
 KW Addison's disease; adrenal gland hyperfunction; obesity.
 XX

OS Bos taurus.
 XX

PN WO200135984-A1.
 XX

PD 25-MAY-2001.
 XX

XX 17-NOV-2000; 2000WO-JP08119.
 XX

XX 18-NOV-1999; 99JP-0327900.
 PR

PR 26-SEP-2000; 2000JP-0297073.
 XX

XX (TAKE) TAKEDA CHEM IND LTD.
 PA

XX Kitada C, Matsumoto H, Hinuma S;
 PI

XX WPI; 2001-355552/37.
 DR

XX Use of G protein receptor ligand or peptide for controlling
 PT corticotropin releasing hormone secretion -

XX Disclosure; Page 64; 90pp; Japanese.
 PS

XX The present sequence describes a method of controlling the secretion of
 CC corticotrophin releasing hormone (CRH), involving the use of a G protein
 CC receptor ligand. This can be used to control the secretion of CRH and is
 CC useful as an analgesic or for treating, preventing or ameliorating
 CC diseases associated with CRH secretion such as hyperaldosteronism,
 CC hypercortisolaemia, secondary or chronic hypoadrenocorticism, Addison's
 CC disease (including boredom, nausea, pigmentation, hypogonadism, hair
 CC loss, and hypotension), adrenal gland hypofunction and obesity. The
 CC present sequence is a peptide used in the exemplification of the
 CC invention.
 XX

SQ Sequence 32 AA;

Query Match 100.0%; Score 172; DB 22; Length 32;
 Best Local Similarity 100.0%; Pred. NO. 1.2e-18;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRAHQSHMEIRTPDINPAWYAGRGIRPVGRF 31
 |||||
 Db 1 SRAHQSHMEIRTPDINPAWYAGRGIRPVGRF 31

Search completed: December 3, 2003, 19:11:24
 Job time : 34.75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 19:09:05 ; Search time 11.5 seconds
(without alignments)
114.055 Million cell updates/sec

Title: US-09-868-885B-3

Perfect score: 172
Sequence: 1 SRAHQHMEIRTPDINPAWYAGRIPVGRF 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	172	100.0	31	3	US-09-105-678A-7
2	172	100.0	31	3	US-09-105-678A-31
3	172	100.0	31	3	US-08-776-971-5
4	172	100.0	31	3	US-08-776-971-97
5	172	100.0	31	3	US-09-421-208-7
6	172	100.0	31	3	US-09-421-208-31
7	172	100.0	31	4	US-09-560-915-13
8	172	100.0	32	3	US-09-105-678A-32
9	172	100.0	32	3	US-08-776-971-6
10	172	100.0	32	3	US-09-421-208-32
11	172	100.0	32	3	US-09-105-678A-33
12	172	100.0	33	3	US-08-776-971-7
13	172	100.0	33	3	US-09-421-208-33
14	172	100.0	98	3	US-08-776-971-1
15	172	100.0	98	3	US-08-776-971-44
16	172	100.0	98	3	US-08-776-971-122
17	172	100.0	98	3	US-08-776-971-131
18	172	100.0	98	3	US-08-776-971-136
19	168	97.7	98	3	US-08-776-971-115
20	168	97.7	98	3	US-08-776-971-117
21	163	94.8	31	3	US-09-105-678A-8
22	163	94.8	31	3	US-09-105-678A-37
23	163	94.8	31	3	US-09-172-353-4
24	163	94.8	31	3	US-08-776-971-47
25	163	94.8	31	3	US-09-421-208-8
26	163	94.8	31	3	US-09-421-208-37
27	163	94.8	31	4	US-09-560-915-14

163 94.8 31 4 US-09-799-955-4 Sequence 4, Appl
29 163 94.8 32 3 US-09-105-678A-38 Sequence 38, Appl
30 163 94.8 32 3 US-08-776-971-48 Sequence 48, Appl
31 163 94.8 32 3 US-09-421-208-38 Sequence 38, Appl
32 163 94.8 33 3 US-09-105-678A-39 Sequence 39, Appl
33 163 94.8 33 3 US-08-776-971-49 Sequence 49, Appl
34 163 94.8 33 3 US-09-421-208-39 Sequence 39, Appl
35 163 94.8 83 3 US-08-776-971-45 Sequence 45, Appl
36 163 94.8 83 3 US-08-776-971-124 Sequence 124, App
37 163 94.8 83 3 US-08-776-971-137 Sequence 137, App
38 161 93.6 29 3 US-09-105-678A-29 Sequence 29, Appl
39 161 93.6 29 3 US-08-776-971-3 Sequence 3, Appl
40 161 93.6 29 3 US-09-421-208-29 Sequence 29, Appl
41 158 91.9 31 3 US-09-105-678A-9 Sequence 9, Appl
42 158 91.9 31 3 US-09-105-678A-43 Sequence 43, Appl
43 158 91.9 31 3 US-08-776-971-61 Sequence 61, Appl
44 158 91.9 31 3 US-09-421-208-9 Sequence 9, Appl
45 158 91.9 31 3 US-09-421-208-43 Sequence 43, Appl

ALIGNMENTS

RESULT 1
US-09-105-678A-7
; Sequence 7, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,678A
; FILING DATE: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 1721118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-105-678A-7

Query Match 100.0%; Score 172; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.3e-19;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SRAHQHMEIRTPDINPAWYAGRIPVGRF 31

Db 1 SRAHQSMSEIRTPDINPANYAGRGIRPVGRF 31
|||||
RESULT 2
US-09-105-678A-31
; Sequence 31, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105.678A
; FILING DATE: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; STRANDEDNESS: linear
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-105-678A-31
Query Match 100.0%; Score 172; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.3e-19;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SRAHQSMSEIRTPDINPANYAGRGIRPVGRF 31
|||||
Db 1 SRAHQSMSEIRTPDINPANYAGRGIRPVGRF 31
|||||
RESULT 3
US-08-776-971-5
; Sequence 5, Application US/08776971B
; Patent No. 6228984
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; Habata, Yugo
; Kawamata, Yuji
; Hosoya, Masaki
; Fujii, Ryo
; Fukusumi, Shoji
; Kitada, Chieko
; TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
; NUMBER OF SEQUENCES: 140
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/776.971B
; FILING DATE: 06-Feb-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP96/03821
; FILING DATE: 28-DEC-1996
; APPLICATION NUMBER: JP 7/343371
; FILING DATE: 28-DEC-1995
; APPLICATION NUMBER: JP 8/59419
; FILING DATE: 15-MAR-1996
; APPLICATION NUMBER: JP 8/211805
; FILING DATE: 12-AUG-1996
; APPLICATION NUMBER: JP 8/246573
; FILING DATE: 18-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 47176
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-08-776-971-5
Query Match 100.0%; Score 172; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.3e-19;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SRAHQSMSEIRTPDINPANYAGRGIRPVGRF 31
|||||
Db 1 SRAHQSMSEIRTPDINPANYAGRGIRPVGRF 31
|||||
RESULT 4
US-08-776-971-97
; Sequence 97, Application US/08776971B
; Patent No. 6228984
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; Habata, Yugo
; Kawamata, Yuji
; Hosoya, Masaki
; Fujii, Ryo
; Fukusumi, Shoji
; Kitada, Chieko
; TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
; NUMBER OF SEQUENCES: 140
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/776.971B
; FILING DATE: 06-Feb-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP96/03821
; FILING DATE: 28-DEC-1996
; APPLICATION NUMBER: JP 7/343371
; FILING DATE: 28-DEC-1995
; APPLICATION NUMBER: JP 8/59419
; FILING DATE: 15-MAR-1996
; APPLICATION NUMBER: JP 8/211805
; FILING DATE: 12-AUG-1996
; APPLICATION NUMBER: JP 8/246573
; FILING DATE: 18-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 47176
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-08-776-971-5
Query Match 100.0%; Score 172; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.3e-19;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SRAHQSMSEIRTPDINPANYAGRGIRPVGRF 31
|||||
Db 1 SRAHQSMSEIRTPDINPANYAGRGIRPVGRF 31
|||||

COMPUTER READABLE FORM:
ZIP: 02109
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/776,971B
FILING DATE: 06-Feb-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP96/03821
FILING DATE: 28-DEC-1996
APPLICATION NUMBER: JP 7/343371
FILING DATE: 28-DEC-1995
APPLICATION NUMBER: JP 8/59419
FILING DATE: 15-MAR-1996
APPLICATION NUMBER: JP 8/211805
FILING DATE: 12-AUG-1996
APPLICATION NUMBER: JP 8/246573
FILING DATE: 18-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 47176
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 97:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 97:
US-08-776-971-97
Query Match 100.0%; Score 172; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.3e-19;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SRAHQHSMETPTDINPAWYAGRGIRPVGRF 31
DB 1 SRAHQHSMETPTDINPAWYAGRGIRPVGRF 31
RESULT 5
US-09-421-208-7
Sequence 7, Application US/09421208
Patent No. 6258561
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/421,208
FILING DATE: 26-JUN-1998
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS: linear
TOPOLOGY: linear

APPLICATION NUMBER: US 09/105,678
FILING DATE: 26-JUN-1998
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS: linear
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-421-208-7
Query Match 100.0%; Score 172; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.3e-19;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SRAHQHSMETPTDINPAWYAGRGIRPVGRF 31
DB 1 SRAHQHSMETPTDINPAWYAGRGIRPVGRF 31
RESULT 6
US-09-421-208-31
Sequence 31, Application US/09421208
Patent No. 6258561
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/105,678
FILING DATE: 26-JUN-1998
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS: linear
TOPOLOGY: linear

```

; MOLECULE TYPE: peptide
US-09-421-208-31
Query Match      100.0%; Score 172; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.3e-19;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSEIRTPDINPAWAGRGIRPVGRF 31
Db 1 SRAHQHSEIRTPDINPAWAGRGIRPVGRF 31

RESULT 7
US-09-560-915-13
; Sequence 13, Application US/09560915
; Patent No. 6383764
; GENERAL INFORMATION:
; APPLICANT: Civeilli, Olivier
; APPLICANT: Lin, Steven
; TITLE OF INVENTION: Relating To Prolactin Releasing Peptide (PrP)
; FILE REFERENCE: P-UC 3534
; CURRENT APPLICATION NUMBER: US/09/560,915
; CURRENT FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 31
; TYPE: PrT
; ORGANISM: Bos taurus
US-09-560-915-13

Query Match      100.0%; Score 172; DB 4; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.3e-19;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSEIRTPDINPAWAGRGIRPVGRF 31
Db 1 SRAHQHSEIRTPDINPAWAGRGIRPVGRF 31

RESULT 8
US-09-105-678A-32
; Sequence 32, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105.678A
; FILING DATE: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026

; MOLECULE TYPE: peptide
US-09-421-208-31
Query Match      100.0%; Score 172; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.3e-19;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSEIRTPDINPAWAGRGIRPVGRF 31
Db 1 SRAHQHSEIRTPDINPAWAGRGIRPVGRF 31

RESULT 9
US-08-776-971-6
; Sequence 6, Application US/08776971B
; Patent No. 6238984
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Habata, Yugo
; APPLICANT: Kawamata, Yuji
; APPLICANT: Hosoya, Masaki
; APPLICANT: Fujii, Ryo
; APPLICANT: Fukusumi, Shoji
; APPLICANT: Kitada, Chieko
; TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
; NUMBER OF SEQUENCES: 140
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/776,971B
; FILING DATE: 06-Feb-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP96/03821
; FILING DATE: 28-DEC-1996
; APPLICATION NUMBER: JP 7/343371
; FILING DATE: 28-DEC-1995
; APPLICATION NUMBER: JP 8/59419
; FILING DATE: 15-MAR-1996
; APPLICATION NUMBER: JP 8/211805
; FILING DATE: 12-AUG-1996
; APPLICATION NUMBER: JP 8/246573
; FILING DATE: 18-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 47176
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 amino acids
```



```
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ FRAGMENT TYPE: internal
/ SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-08-776-971-6
Query Match 100.0%; Score 172; DB 3; Length 32;
Best Local Similarity 100.0%; Pred. No. 1.3e-19;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SRAHQHSEIRTPDINPAWYAGRGIRPVGRF 31
DB 1 SRAHQHSEIRTPDINPAWYAGRGIRPVGRF 31

RESULT 10
US-09-421-208-32
Sequence 32, Application US/09421208
Patent No. 6258561
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/421,208
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/105,678
FILING DATE: 26-JUN-1998
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 32 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-421-208-32
Query Match 100.0%; Score 172; DB 3; Length 32;
Best Local Similarity 100.0%; Pred. No. 1.3e-19;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SRAHQHSEIRTPDINPAWYAGRGIRPVGRF 31
DB 1 SRAHQHSEIRTPDINPAWYAGRGIRPVGRF 31

RESULT 11
US-09-105-678A-33
Sequence 33, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-33
Query Match 100.0%; Score 172; DB 3; Length 33;
Best Local Similarity 100.0%; Pred. No. 1.4e-19;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SRAHQHSEIRTPDINPAWYAGRGIRPVGRF 31
DB 1 SRAHQHSEIRTPDINPAWYAGRGIRPVGRF 31

RESULT 12
US-08-776-971-7
Sequence 7, Application US/08776971B
Patent No. 6228984
GENERAL INFORMATION:
APPLICANT: Hinuma, Shuji
APPLICANT: Habata, Yugo
APPLICANT: Kawamata, Yuji
APPLICANT: Hosoya, Masaki
APPLICANT: Fujii, Ryo
APPLICANT: Fukusumi, Shoji
APPLICANT: Kitada, Chieko
TITLE OF INVENTION: POLYPEPTIDES, THEIR PRODUCTION AND USE
NUMBER OF SEQUENCES: 140
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
```



```
/
/
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Conlin, David G.
/ REGISTRATION NUMBER: 27,026
/ REFERENCE/DOCKET NUMBER: 47176
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617-523-3400
/ TELEFAX: 617-523-6440
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 98 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ FRAGMENT TYPE: internal
/ SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-08-776-971-1

Query Match          100.0%; Score 172; DB 3; Length 98;
Best Local Similarity 100.0%; Pred. No. 5.1e-19;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSMETPTDINPANYAGRGIRPVGRF 31
Db 23 SRAHQHSMETPTDINPANYAGRGIRPVGRF 53

RESULT 15
US-08-776-971-44
; Sequence 44, Application US/08776971B
; Patent No. 6228984
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; Habata, Yugo
; Kawamata, Yuji
; Hosoya, Masaki
; Fujii, Ryo
; Fukusumi, Shoji
; Kitada, Chieko
; TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
; NUMBER OF SEQUENCES: 140
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/776,971B
; FILING DATE: 06-Feb-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP96/03821
; FILING DATE: 28-DEC-1996
; APPLICATION NUMBER: JP 7/343371
; FILING DATE: 28-DEC-1995
; APPLICATION NUMBER: JP 8/59419
; FILING DATE: 15-MAR-1996
; APPLICATION NUMBER: JP 8/211805
; FILING DATE: 12-AUG-1996
; APPLICATION NUMBER: JP 8/246573
; FILING DATE: 18-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 47176
; TELECOMMUNICATION INFORMATION:
/
/
/ TELEPHONE: 617-523-3400
/ TELEFAX: 617-523-6440
/ INFORMATION FOR SEQ ID NO: 44:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 98 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ FRAGMENT TYPE: internal
/ SEQUENCE DESCRIPTION: SEQ ID NO: 44:
US-08-776-971-44

Query Match          100.0%; Score 172; DB 3; Length 98;
Best Local Similarity 100.0%; Pred. No. 5.1e-19;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSMETPTDINPANYAGRGIRPVGRF 31
Db 23 SRAHQHSMETPTDINPANYAGRGIRPVGRF 53

Search completed: December 3, 2003, 19:15:55
Job time : 12.5 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 19:11:30 ; Search time 22.5 Seconds
(without alignments)
256.244 Million cell updates/sec

Title: US-09-868-885B-3

Perfect score: 172

Sequence: 1 SRAHQSMETPTDINPAWYAGRGIRPVGRF 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpa/PTCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpa/PTCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpa/US09D_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	172	100.0	31	9	US-09-932-161-13
2	172	100.0	31	12	US-10-096-777-13
3	172	100.0	31	14	US-10-044-592-39
4	172	100.0	32	14	US-10-044-592-40
5	172	100.0	33	14	US-10-044-592-41
6	172	100.0	98	14	US-10-044-592-28
7	172	100.0	98	14	US-10-044-592-38
8	172	100.0	98	14	US-10-044-592-82
9	172	100.0	98	14	US-10-044-592-84
10	172	100.0	98	14	US-10-044-592-86
11	172	100.0	98	14	US-10-044-592-88
12	163	94.8	31	9	US-09-932-161-14
13	163	94.8	31	12	US-10-096-777-14
14	163	94.8	31	14	US-10-044-592-4
15	163	94.8	31	14	US-10-044-592-5

16	163	94.8	70	14	US-10-044-592-90	Sequence 90, Appl
17	163	94.8	82	14	US-10-044-592-1	Sequence 1, Appl
18	163	94.8	86	14	US-10-044-592-96	Sequence 96, Appl
19	163	94.8	91	14	US-10-044-592-94	Sequence 94, Appl
20	161	93.6	29	14	US-10-044-592-26	Sequence 26, Appl
21	158	91.9	31	9	US-09-932-161-15	Sequence 15, Appl
22	158	91.9	31	12	US-10-096-777-15	Sequence 15, Appl
23	158	91.9	87	14	US-10-044-592-92	Sequence 92, Appl
24	142	82.6	25	14	US-10-044-592-78	Sequence 78, Appl
25	115	66.9	20	9	US-09-932-161-16	Sequence 16, Appl
26	115	66.9	20	12	US-10-096-777-16	Sequence 16, Appl
27	115	66.9	20	14	US-10-044-592-42	Sequence 42, Appl
28	115	66.9	21	14	US-10-044-592-43	Sequence 43, Appl
29	115	66.9	22	14	US-10-044-592-44	Sequence 44, Appl
30	111	64.5	20	9	US-09-932-161-17	Sequence 17, Appl
31	111	64.5	20	12	US-10-096-777-17	Sequence 17, Appl
32	111	64.5	20	14	US-10-044-592-6	Sequence 6, Appl
33	109	63.4	19	14	US-10-044-592-27	Sequence 27, Appl
34	109	63.4	20	9	US-09-932-161-18	Sequence 18, Appl
35	109	63.4	20	12	US-10-096-777-18	Sequence 18, Appl
36	96	55.8	40	14	US-10-044-592-80	Sequence 80, Appl
37	61	35.5	428	9	US-09-820-155-2	Sequence 2, Appl
38	57	33.1	9	14	US-10-044-592-8	Sequence 8, Appl
39	57	33.1	465	12	US-10-301-822-197	Sequence 197, App
40	56	32.6	428	9	US-09-820-155-4	Sequence 4, Appl
41	54	31.4	10	14	US-10-044-592-9	Sequence 9, Appl
42	54	31.4	209	14	US-10-108-915-30	Sequence 30, Appl
43	52	30.2	417	10	US-09-738-626-4642	Sequence 4642, Ap
44	52	30.2	3594	10	US-09-911-842-4	Sequence 4, Appl
45	52	30.2	3594	14	US-10-150-821-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-09-932-161-13
; Sequence 13, Application US/09932161
; Patent No. US20020037533A1
; GENERAL INFORMATION:
; APPLICANT: Civeilli, Olivier
; APPLICANT: Lin, Steven
; TITLE OF INVENTION: Screening and Therapeutic Methods For
; TITLE OF INVENTION: Promoting Wakefulness and Sleep
; FILE REFERENCE: P-UC 4679
; CURRENT APPLICATION NUMBER: US/09/932,161
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 09/560,915
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Bos taurus
US-09-932-161-13

Query Match 100.0%; Score 172; DB 9; Length 31;
Best Local Similarity 100.0%; Pred. No. 8.2e-18;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRAHQSMETPTDINPAWYAGRGIRPVGRF 31
Db 1 SRAHQSMETPTDINPAWYAGRGIRPVGRF 31

RESULT 2
US-10-096-777-13
; Sequence 13, Application US/10096777
; Publication No. US20030171270A1
; GENERAL INFORMATION:
; APPLICANT: Civeilli, Olivier
; APPLICANT: Lin, Steven

```
; TITLE OF INVENTION: Therapeutic Compositions and Methods
; TITLE OF INVENTION: Relating To Prolactin Releasing Peptide (PrRP)
; FILE REFERENCE: P-UC 3534
; CURRENT APPLICATION NUMBER: US/10/096,777
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: US/09/560,915
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Bos taurus
US-10-096-777-13

Query Match      100.0%; Score 172; DB 12; Length 31;
Best Local Similarity 100.0%; Pred. No. 8.2e-18;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSMETPDINPAWYAGRGIRPVGRF 31
Db 1 SRAHQHSMETPDINPAWYAGRGIRPVGRF 31

RESULT 3
US-10-044-592-39
; Sequence 39, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 39
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-39

Query Match      100.0%; Score 172; DB 14; Length 31;
Best Local Similarity 100.0%; Pred. No. 8.2e-18;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSMETPDINPAWYAGRGIRPVGRF 31
Db 1 SRAHQHSMETPDINPAWYAGRGIRPVGRF 31

RESULT 4
US-10-044-592-40
; Sequence 40, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 40
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-40

Query Match      100.0%; Score 172; DB 12; Length 31;
Best Local Similarity 100.0%; Pred. No. 8.2e-18;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSMETPDINPAWYAGRGIRPVGRF 31
Db 1 SRAHQHSMETPDINPAWYAGRGIRPVGRF 31

RESULT 5
US-10-044-592-41
; Sequence 41, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 41
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-41

Query Match      100.0%; Score 172; DB 14; Length 33;
Best Local Similarity 100.0%; Pred. No. 8.7e-18;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSMETPDINPAWYAGRGIRPVGRF 31
Db 1 SRAHQHSMETPDINPAWYAGRGIRPVGRF 31

RESULT 6
US-10-044-592-28
; Sequence 28, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 28
```

```
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Murine
US-10-044-592-28

Query Match      100.0%; Score 172; DB 14; Length 98;
Best Local Similarity 100.0%; Pred. No. 2.7e-17;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSEIIRTPDINPAWYAGRGIRPVGRF 31
DB 23 SRAHQHSEIIRTPDINPAWYAGRGIRPVGRF 53

RESULT 7
US-10-044-592-38
; Sequence 38, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Fukusumi, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 38
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-38

Query Match      100.0%; Score 172; DB 14; Length 98;
Best Local Similarity 100.0%; Pred. No. 2.7e-17;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSEIIRTPDINPAWYAGRGIRPVGRF 31
DB 23 SRAHQHSEIIRTPDINPAWYAGRGIRPVGRF 53

RESULT 8
US-10-044-592-82
; Sequence 82, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Fukusumi, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 82
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-82

Query Match      100.0%; Score 172; DB 14; Length 98;
Best Local Similarity 100.0%; Pred. No. 2.7e-17;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSEIIRTPDINPAWYAGRGIRPVGRF 31
DB 23 SRAHQHSEIIRTPDINPAWYAGRGIRPVGRF 53

RESULT 9
US-10-044-592-84
; Sequence 84, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Fukusumi, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 84
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-84

Query Match      100.0%; Score 172; DB 14; Length 98;
Best Local Similarity 100.0%; Pred. No. 2.7e-17;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSEIIRTPDINPAWYAGRGIRPVGRF 31
DB 23 SRAHQHSEIIRTPDINPAWYAGRGIRPVGRF 53

RESULT 10
US-10-044-592-86
; Sequence 86, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Fukusumi, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 86
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-86

Query Match      100.0%; Score 172; DB 14; Length 98;
Best Local Similarity 100.0%; Pred. No. 2.7e-17;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSEIIRTPDINPAWYAGRGIRPVGRF 31
DB 23 SRAHQHSEIIRTPDINPAWYAGRGIRPVGRF 53
```

```
Db          23 SRAHQHSMETPTDINPAWYAGRGIRPVGRF 53
|||||
RESULT 11
US-10-044-592-88
; Sequence 88, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 88
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-88
Query Match          100.0%; Score 172; DB 14; Length 98;
Best Local Similarity 100.0%; Pred. No. 2.7e-17;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy          1 SRAHQHSMETPTDINPAWYAGRGIRPVGRF 31
|||||
Db          23 SRAHQHSMETPTDINPAWYAGRGIRPVGRF 53
|||||
RESULT 12
US-09-932-161-14
; Sequence 14, Application US/09932161
; Patent No. US20020037533A1
; GENERAL INFORMATION:
; APPLICANT: Civelii, Olivier
; TITLE OF INVENTION: Screening and Therapeutic Methods For
; FILE REFERENCE: P-UC 4679
; CURRENT APPLICATION NUMBER: US/09/932,161
; CURRENT FILING DATE: 2001-08-17
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Rattus
US-09-932-161-14
Query Match          94.8%; Score 163; DB 9; Length 31;
Best Local Similarity 93.5%; Pred. No. 1.6e-16;
Matches 29; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy          1 SRAHQHSMETPTDINPAWYAGRGIRPVGRF 31
|||||
Db          1 SRAHQHSMETPTDINPAWYAGRGIRPVGRF 31
|||||
RESULT 13
US-10-096-777-14
; Sequence 14, Application US/10096777
; Publication No. US2003017270A1
; GENERAL INFORMATION:
; APPLICANT: Civelii, Olivier
; TITLE OF INVENTION: Therapeutic Compositions and Methods
; FILE REFERENCE: P-UC 3534
; CURRENT APPLICATION NUMBER: US/10/096,777
; CURRENT FILING DATE: 2002-03-12
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Rattus
US-10-096-777-14
Query Match          94.8%; Score 163; DB 12; Length 31;
Best Local Similarity 93.5%; Pred. No. 1.6e-16;
Matches 29; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy          1 SRAHQHSMETPTDINPAWYAGRGIRPVGRF 31
|||||
Db          1 SRAHQHSMETPTDINPAWYAGRGIRPVGRF 31
|||||
RESULT 14
US-10-044-592-4
; Sequence 4, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 4
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Murine
US-10-044-592-4
Query Match          94.8%; Score 163; DB 14; Length 31;
Best Local Similarity 93.5%; Pred. No. 1.6e-16;
Matches 29; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy          1 SRAHQHSMETPTDINPAWYAGRGIRPVGRF 31
|||||
Db          1 SRAHQHSMETPTDINPAWYAGRGIRPVGRF 31
|||||
RESULT 15
US-10-044-592-5
; Sequence 5, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
```

```
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 5
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(31)
; OTHER INFORMATION: antigen
US-10-044-592-5

Query Match      94.8%; Score 163; DB 14; Length 31;
Best Local Similarity 93.5%; Pred. No. 1.6e-16;
Matches 29; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 SRAHSHMEIRTPDINPAWYAGRGIRPVGRF 31
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      1 SRAHSHMETRTPDINPAWYTCRGIRPVGRF 31
```

Search completed: December 3, 2003, 19:17:37
Job time : 22.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

QM protein - protein search, using sw model

Run on: December 3, 2003, 19:05:30 ; Search time 10.5 seconds
(without alignment)
283.927 Million cell updates/sec

Title: US-09-868-885B-3
Perfect score: 172
Sequence: 1 SRAHQSHWEIRTPDINPAWAGRGIRPVGRF 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76:*
1: PIR1.*
2: PIR2.*
3: PIR3.*
4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	163	94.8	83	2 JC7607	prolactin-releasin
2	56.5	32.8	303	2 AH2016	hypothetical prote
3	55	32.0	1236	2 T50904	Mg protoporphyrin
4	54.5	31.7	503	2 A82193	Sun/nucleolar prot
5	53.5	31.1	176	2 S87150	hypothetical prote
6	53	30.8	376	2 C75580	adenine deaminase-
7	52	30.2	73	2 A98137	hypothetical prote
8	52	30.2	308	2 AH2896	transcription regu
9	52	30.2	308	2 B97672	rok family protein
10	52	30.2	1292	2 AH3166	hypothetical prote
11	52	30.2	419	2 T31462	probable magnesium
12	51.5	29.9	664	2 F83376	conserved hypothet
13	51	29.7	294	2 T21075	hypothetical prote
14	51	29.7	430	1 B69009	conserved hypothet
15	51	29.7	798	2 S11210	probable unr prote
16	51	29.7	962	2 H89157	exonuclease ABC c
17	50.5	29.4	790	2 T47959	hypothetical prote
18	50	29.1	527	2 T33175	hypothetical prote
19	50	29.1	974	2 AH3361	exonuclease ABC c
20	49.5	28.8	393	2 AB2664	conserved hypothet
21	49.5	28.8	401	2 A97446	hypothetical prote
22	49	28.5	128	2 S76955	hypothetical prote
23	49	28.5	226	2 A87664	hypothetical prote
24	48.5	28.2	254	2 S76814	hypothetical prote
25	48.5	28.2	548	2 T47548	hypothetical prote
26	48.5	28.2	1882	2 S73484	hypothetical prote
27	48	27.9	72	2 E91002	probable regulator
28	48	27.9	220	2 C83292	probable glucathio
29	48	27.9	314	2 B70569	hypothetical prote

30	48	27.9	348	2 T21648	hypothetical prote
31	48	27.9	365	2 T39098	hypothetical prote
32	48	27.9	397	2 JC7675	acetoacetyl-CoA re
33	48	27.9	424	2 B38176	samb protein - Sal
34	48	27.9	637	2 D87559	sensory box histid
35	48	27.9	772	2 T07958	protoporphylin IX
36	48	27.9	960	2 A71315	exonuclease ABC c
37	48	27.9	973	2 AH2762	ABC exonuclease g
38	48	27.9	982	2 G97543	exonuclease ABC c
39	48	27.9	1328	2 AE2351	protoporphylin IX
40	48	27.9	1331	2 S75000	protoporphylin IX
41	48	27.9	1379	2 S37310	protoporphylin IX
42	48	27.9	1380	2 S64721	protoporphylin IX
43	48	27.9	1381	2 S71388	protoporphylin IX
44	48	27.9	1382	2 T01789	protoporphylin IX
45	48	27.9	1383	2 T07126	magnesium chelatas

ALIGNMENTS

RESULT 1

JC7607
prolactin-releasing peptide - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 30-Jun-2001
C:Accession: JC7607
R;Yamada, M.; Ozawa, A.; Iehii, S.; Shibusawa, N.; Hashida, T.; Hosoya, T.
Biochem. Biophys. Res. Commun. 281, 53-56, 2001
A:Title: Isolation and characterization of the rat prolactin-releasing peptide gene: Mu
A:Reference number: JC7607; MUID:21092785; PMID:11178959
A:Contents: Spleen
A:Accession: JC7607
A:Molecule type: DNA
A:Residues: 1-83 <YAM>
A:Cross-references: DDBJ:AB040612; DDBJ:AB040613
C:Comment: This peptide induces arachidonic acid metabolite release from rat anterior p
release, and stimulation of ACTH secretion from the pituitary.
C:Genetics:
A:Gene: PrRP
A:Introns: 33/1

Query Match 94.8%; Score 163; DB 2; Length 83;
Best Local Similarity 93.5%; Pred. No. 6.9e-17;
Matches 29; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 SRAHQSHWEIRTPDINPAWAGRGIRPVGRF 31
|||||
Db 22 SRAHQSHWEIRTPDINPAWAGRGIRPVGRF 52
|||||

RESULT 2

AH2016
hypothetical protein all1686 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C:Accession: AH2016
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasaamoto, S.; Watanabe, A.; Iriguch
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, .
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AH2016
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-303 <KUR>
A:Cross-references: GB:BA000019; FIDN:BA878052.1; FID:gl17135506; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all1686

Query Match 32.8%; Score 56.5; DB 2; Length 303;

Best Local Similarity 51.6%; Pred. No. 1.5;
Matches 16; Conservative 2; Mismatches 8; Indels 5; Gaps 2;

QY 4 HOHSMSEIRTPDINPAWY-----AGRGIRPVGR 30
DB 226 HEHSYE-RTAIDGTYYLTCTGAGAGNRPVGR 255

RESULT 3
T50904
Mg protoporphyrin methyl transferase [imported] - Rubrivivax gelatinosus
C;Species: Rubrivivax gelatinosus
C;Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 02-Sep-2000
C;Accession: T50904
R;Nagashima, K.V.; Igarashi, N.; Harada, J.; Nagasahima, S.; Matsura, K.; Shimada, K.
submitted to the EMBL Data Library, November 1999
A;Description: Determination of Nucleotide Sequences of Rubrivivax gelatinosus Photosyn
A;Reference number: Z25270
A;Accession: T50904
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1236 <NAG>
A;Cross-references: EMBL:AB034704; PIDN:BAA94057.1
A;Experimental source: strain IL144
C;Genetics:
A;Gene: bchH
C;Superfamily: Rhodobacter capsulatus magnesium-protoporphyrin O-methyltransferase

Query Match 32.0%; Score 55; DB 2; Length 1236;
Best Local Similarity 34.4%; Pred. No. 11;
Matches 11; Conservative 6; Mismatches 9; Indels 6; Gaps 1;

QY 3 AHQHSMSEIRTPDINPAWYAG-----RGIRPV 28
DB 1112 SEQVALETRTMLNPKNWEGMLHGVEGVRI 1143

RESULT 4
A82193
Sun/nucleolar protein family protein VCI1502 [imported] - Vibrio cholerae (strain N16961
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C;Accession: A82193
R;Heidelbergl, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
Chardson, D.; Ermolaeva, M.D.; Vanathavan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, B.
I., R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; UID:20406833; PMID:10952301
A;Accession: A82193
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-503 <HBI>
A;Cross-references: GB:AE004428; GB:AE003852; NID:g96555997; PIDN:AAF94657.1; GSPDB:GN001
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VCI502
A;Map position: 1

Query Match 31.7%; Score 54.5; DB 2; Length 503;
Best Local Similarity 36.7%; Pred. No. 5;
Matches 11; Conservative 6; Mismatches 12; Indels 1; Gaps 1;

QY 1 SRAHQHSMSEIRTPDINPAWYAGRGIRPVGR 30
DB 435 SSSASHSVELDTTOAR-EWFMGDRVRPEQG 463

RESULT 5
S67150
hypothetical protein YOR253w - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein O5315
C;Species: Saccharomyces cerevisiae


~~~~~

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 19:01:34 ; Search time 6.25 Seconds  
(without alignments)  
233.252 Million cell updates/sec

Title: US-09-868-885B-3  
Perfect score: 172  
Sequence: 1 SRAHQHSMETIRPDINPAWYAGRGIRPVGRF 31

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID         | Description        |
|------------|-------|-------------|--------|---------------|--------------------|
| 1          | 172   | 100.0       | 98     | 1 PRRP_BOVIN  | P81264 bos taurus  |
| 2          | 163   | 94.8        | 83     | 1 PRRP_RAT    | P81278 rattus norv |
| 3          | 158   | 91.9        | 87     | 1 PRRP_HUMAN  | P81277 homo sapien |
| 4          | 61    | 35.5        | 428    | 1 NER3_BOVIN  | Q97859 bos taurus  |
| 5          | 56    | 32.6        | 428    | 1 NER3_HUMAN  | Q9uq49 homo sapien |
| 6          | 52    | 30.2        | 417    | 1 EX7L_CORGL  | Q9ntm3 corynebacte |
| 7          | 52    | 30.2        | 447    | 1 EX7L_COREP  | Q8fgp1 corynebacte |
| 8          | 51    | 29.7        | 798    | 1 UNR_RAT     | P18395 rattus norv |
| 9          | 51    | 29.7        | 962    | 1 UVR_A_METH  | Q26543 methanobact |
| 10         | 50    | 29.1        | 288    | 1 Y587_PASMU  | Q9cm56 pasteurella |
| 11         | 50    | 29.1        | 973    | 1 UVR_A_RHIL  | Q98m36 rhizobium l |
| 12         | 50    | 29.1        | 973    | 1 UVR_A_RHIME | P56899 rhizobium m |
| 13         | 48.5  | 28.2        | 239    | 1 6PGL_SYNY3  | P74618 synchocyst  |
| 14         | 48.5  | 28.2        | 1882   | 1 Y468_MYCPN  | P75109 mycoplasma  |
| 15         | 48    | 27.9        | 424    | 1 SAMB_SALTY  | P23832 salmonella  |
| 16         | 48    | 27.9        | 960    | 1 UVR_A_TREPA | Q83527 treponema p |
| 17         | 47.5  | 27.6        | 1083   | 1 T2D3_HUMAN  | Q00288 homo sapien |
| 18         | 47    | 27.3        | 406    | 1 DH12_RABIT  | P51976 oryctolagus |
| 19         | 47    | 27.3        | 418    | 1 NER3_RAT    | Q99pw5 rattus norv |
| 20         | 47    | 27.3        | 442    | 1 TRB2_METMA  | Q8q001 methanosarc |
| 21         | 47    | 27.3        | 940    | 1 UVR_A_VIBCH | Q9kuw5 vibrio chol |
| 22         | 47    | 27.3        | 969    | 1 UVR_A_MYCLE | Q9cc24 mycobacteri |
| 23         | 47    | 27.3        | 972    | 1 UVR_A_MYCTU | P94972 mycobacteri |
| 24         | 47    | 27.3        | 1308   | 1 M4K6_MOUSE  | Q9jms2 mus musculu |
| 25         | 46.5  | 27.0        | 345    | 1 ARG6_BACHD  | Q9k8v2 bacillus ha |
| 26         | 46.5  | 27.0        | 770    | 1 AVP3_ARATH  | P31414 arabidopsis |
| 27         | 46    | 26.7        | 277    | 1 DAPF_CORGL  | Q8np73 corynebacte |
| 28         | 46    | 26.7        | 342    | 1 Y762_METJA  | Q58172 methanococc |
| 29         | 46    | 26.7        | 347    | 1 Y576_METJA  | Q57996 methanococc |
| 30         | 46    | 26.7        | 402    | 1 EX7L_STRCO  | Q9fbm3 streptomyce |
| 31         | 46    | 26.7        | 510    | 1 YCGB_ECOLI  | P29013 escherichia |
| 32         | 46    | 26.7        | 798    | 1 UNR_HUMAN   | Q75534 homo sapien |
| 33         | 46    | 26.7        | 836    | 1 TGM1_RABIT  | P22758 oryctolagus |

34 46 26.7 943 1 UVR\_A\_HAEN  
35 46 26.7 943 1 UVR\_A\_PASMU  
36 45.5 26.5 272 1 TRA2\_DROVI  
37 45.5 26.5 374 1 YRHJ\_SCOLI  
38 45.5 26.5 411 1 PUR7\_ARATH  
39 45.5 26.5 670 1 LIPW\_NEIMB  
40 45.5 26.5 704 1 SUV1\_ARATH  
41 45 26.2 387 1 HIS8\_HAEN1  
42 45 26.2 418 1 NER3\_MOUSE  
43 45 26.2 482 1 R167\_YEAST  
44 45 26.2 676 1 EXL1\_HUMAN  
45 45 26.2 817 1 TGM1\_HUMAN

## ALIGNMENTS

RESULT 1  
PRRP\_BOVIN STANDARD; PRT; 98 AA.  
ID P81264;  
AC 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Prolactin-releasing peptide precursor (PrRP) (Prolactin-releasing hormone) [Contains: Prolactin-releasing peptide PrRP31; Prolactin-releasing peptide PrRP20].  
DE releasing peptide PrRP20].  
GN PRH.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Theria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.  
OC NCBI\_TaxID=9913;  
OX [1]  
RN SEQUENCE FROM N.A., AND SEQUENCE OF 23-52.  
RC TISSUE=Brain;  
RX MEDLINE=98268781; PubMed=9607765;  
RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S., Kitada C., Masuo Y., Asano T., Mateumoto H., Sekiguchi M., Kurokawa T., Nishimura O., Onda H., Fujino M.;  
RT "A prolactin-releasing peptide in the brain.";  
RL Nature 393:272-276(1998).  
CC -1- FUNCTION: Stimulates prolactin (PRL) release and regulates the expression of prolactin through its receptor GPR10. May stimulate lactotrophs directly to secrete PRL.  
CC -1- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC EMBL; AB015417; BAA29025.1; ;  
KW Hormone; Amidation; Signal; Cleavage on pair of basic residues.  
FT SIGNAL 1 22  
FT PEPTIDE 23 53 PROLACTIN-RELEASING PEPTIDE PRRP31.  
FT PEPTIDE 33 53 PROLACTIN-RELEASING PEPTIDE PRRP20.  
FT PROPEP 58 98  
FT MOD\_RES 53 53 AMIDATION (G-54 PROVIDE AMIDE GROUP).  
SQ SEQUENCE 98 AA; 10544 MW; 08AC35A13B0FA908 CRC64;

Query Match 100.0%; Score 172; DB 1; Length 98;  
Best Local Similarity 100.0%; Pred. No. 1.8e-18;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSMETIRPDINPAWYAGRGIRPVGRF 31  
DB 23 SRAHQHSMETIRPDINPAWYAGRGIRPVGRF 53





```

FT ACT_SITE 45 45 Potential.
FT ACT_SITE 50 50 Potential.
FT ACT_SITE 87 87 Potential.
FT ACT_SITE 225 225 Potential.
FT ACT_SITE 245 245 Potential.
FT ACT_SITE 340 340 By similarity.
FT ACT_SITE 370 370 Potential.
FT ACT_SITE 387 387 Potential.
SQ SEQUENCE 428 AA; 48252 MW; 35D1DD9359A78C98 CRC64;

Query Match 32.6%; Score 56; DB 1; Length 428;
Best Local Similarity 37.0%; Pred. No. 1.1;
Matches 10; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 2 RAHQSMETPTDNPANVAGRGIRPV 28
DB 195 KTRPHSLMIYSDDLGVTHHGLIRPM 221

RESULT 6
EX7L_CORGL STANDARD; PRT; 417 AA.
AC Q8NRM3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Probable exodeoxyribonuclease VII large subunit (EC 3.1.11.6)
DE (Exonuclease VII large subunit).
GN XSEA OR CGL1025.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032."
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Bidirectionally degrades single-stranded DNA into large
CC acid-insoluble oligonucleotides, which are then degraded further
CC into small acid-soluble oligonucleotides (By similarity).
CC -1- CATALYTIC ACTIVITY: Exonucleolytic cleavage in either 5'- to 3'-
CC or 3'- to 5'-direction to yield nucleoside 5'-phosphates.
CC -1- SUBUNIT: Heterooligomer composed of large and small subunits (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE XSEA FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AP005277; BAB98418.1; -
CC HAMAP; MF 00378; -; 1.
CC InterPro; IPR003753; Exonuc_VII_L.
CC Pfam; PF02601; Exonuc_VII_L; 1.
CC TIGRFAMs; TIGR00237; xsea; 1.
CC KW Hydrolyase; Nuclease; Exonuclease; Complete proteome.
CC SEQUENCE 417 AA; 45582 MW; B32CD9286C173C34 CRC64;

Query Match 30.2%; Score 52; DB 1; Length 417;
Best Local Similarity 57.1%; Pred. No. 4.1;
Matches 12; Conservative 1; Mismatches 0; Indels 8; Gaps 1;

QY 17 PAWYAGRG-----IRPVG 29
DB 94 PAFYAGRGTFSLWVDIRPVG 114

us-09-868-885b-3.rsp

```

```

RESULT 7
EX7L_COREF STANDARD; PRT; 447 AA.
AC Q8FQPI;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Probable exodeoxyribonuclease VII large subunit (EC 3.1.11.6)
DE (Exonuclease VII large subunit).
GN XSEA OR CE1078.
OS Corynebacterium efficiens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=152794;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;
RA Kawarayashi Y., Yamazaki J., Hino Y., Kikuchi H., Nakamura Y.,
RA Ikeo K., Suzuki M., Mashima J., Itoh T., Yamagishi A., Nishio Y.,
RA Ueda Y., Sugimoto S.;
RT "The entire genomic sequence of Corynebacterium efficiens YS-314."
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Bidirectionally degrades single-stranded DNA into large
CC acid-insoluble oligonucleotides, which are then degraded further
CC into small acid-soluble oligonucleotides (By similarity).
CC -1- CATALYTIC ACTIVITY: Exonucleolytic cleavage in either 5'- to 3'-
CC or 3'- to 5'-direction to yield nucleoside 5'-phosphates.
CC -1- SUBUNIT: Heterooligomer composed of large and small subunits (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE XSEA FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AP005217; BAC17888.1; -
CC HAMAP; MF 00378; -; 1.
CC InterPro; IPR003753; Exonuc_VII_L.
CC Pfam; PF02601; Exonuc_VII_L; 1.
CC TIGRFAMs; TIGR00237; xsea; 1.
CC KW Hydrolyase; Nuclease; Exonuclease; Complete proteome.
CC SEQUENCE 447 AA; 49051 MW; FDBDF6A768E742C1 CRC64;

Query Match 30.2%; Score 52; DB 1; Length 447;
Best Local Similarity 57.1%; Pred. No. 4.4;
Matches 12; Conservative 1; Mismatches 0; Indels 8; Gaps 1;

QY 17 PAWYAGRG-----IRPVG 29
DB 124 PAFYAGRGTFSLWVDIRPVG 144

RESULT 8
UNR_RAT STANDARD; PRT; 798 AA.
AC P18395;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE UNR protein.
GN UNR.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.

```



```

CC TISSUE=Testis;
CC MEDLINE=90370473; PubMed=2204029;
CC Jeffers M., Paciucci R., Pellicer A.;
CC "Characterization of unr; a gene closely linked to N-ras.";
CC Nucleic Acids Res. 18:4891-4899(1990).
CC -1- FUNCTION: RNA-BINDING PROTEIN (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X52311; CAA36549.1; -.
CC PIR; S11210; S11210.
CC HSSP; P15277; IMJC.
CC InterPro; IPR002059; Cold_shock.
CC Pfam; PF00313; CSD; 7.
CC ProDom; PD000621; Cold_shock; 1.
CC SMART; SM00357; CSP; 5.
CC PROSITE; PS00352; COLD_SHOCK; 4.
CC RNA-binding; Repeat.
CC DOMAIN 26 87 CSD 1.
CC FT DOMAIN 136 179 CSD 2 (INCOMPLETE).
CC FT DOMAIN 186 245 CSD 3.
CC FT DOMAIN 297 337 CSD 4 (INCOMPLETE).
CC FT DOMAIN 349 410 CSD 5.
CC FT DOMAIN 447 507 CSD 6.
CC FT DOMAIN 519 579 CSD 7.
CC FT DOMAIN 610 670 CSD 8.
CC FT DOMAIN 674 735 CSD 9.
CC SEQUENCE 798 AA; 88894 MW; F484B3FA8B0995A4 CRC64;
CC
CC Query Match 29.7%; Score 51; DB 1; Length 798;
CC Best Local Similarity 43.5%; Pred. No. 11;
CC Matches 10; Conservative 5; Mismatches 8; Indels 0; Gaps 0;
CC
CC QY 6 HSMWEITPDINAWAGRGIRPV 28
CC Db 583 HSVGITTEANPTIYSGKVRPL 605
CC
CC RESULT 9
CC UVRA METH STANDARD; PRT; 962 AA.
CC AC O26543;
CC DT 30-MAY-2000 (Rel. 39, Created)
CC DT 30-MAY-2000 (Rel. 39, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE UVrABC system protein A (UvrA protein) (Excinuclease ABC subunit A).
CC GN UVRA OR MTH443.
CC OS Methanobacterium thermoautotrophicum.
CC OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
CC OC Methanobacteriaceae; Methanothermobacter.
CC OX NCBI_TaxID=187420;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=Delta H;
CC RX MEDLINE=98037514; PubMed=9371463;
CC RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
CC RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
CC RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
CC RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
CC RA Jiwni N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
CC RA McDougall S., Shiner G., Goyal A., Pietrowski S., Church G.M.,
CC RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
CC RT "Complete genome sequence of Methanobacterium thermoautotrophicum
CC delcA: functional analysis and comparative genomics.";
CC J. Bacteriol. 179:7135-7155(1997).
CC
CC -1- FUNCTION: The UvrABC repair system catalyzes the recognition and
CC processing of DNA lesions. UvrA is an ATPase and a DNA-binding
CC protein. A damage recognition complex composed of 2 uvrA and 2
CC uvrB subunits scans DNA for abnormalities. When the presence of a
CC lesion has been verified by uvrB, the uvrA molecules dissociate
CC (By similarity).
CC -1- SUBUNIT: Forms a heterotrimer with uvrB during the search for
CC lesions (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. UVRA SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE000828; AAB84949.1; -.
CC PIR; H69157; H69157.
CC HAMAP; MP_00205; -.
CC InterPro; IPR003593; AAA ATPase.
CC InterPro; IPR003439; ABC Transporter.
CC InterPro; IPR004602; UvrA.
CC Pfam; PF00005; ABC_tran; 2.
CC ProDom; PD000006; ABC_transporter; 1.
CC SMART; SM00382; AAA; 2.
CC TIGRFAMs; TIGR00630; uvrA; 1.
CC PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
CC PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
CC SOS response; Excision nuclease; DNA repair; DNA recombination;
CC DNA excision; ATP-binding; DNA-binding; Repeat; Zinc; Metal-binding;
CC Zinc-finger; Complete proteome.
CC FT NP_BIND 38 45 ATP (POTENTIAL).
CC FT NP_BIND 649 656 ATP (POTENTIAL).
CC FT ZN_FING 748 774 C4-TYPE
CC SEQUENCE 962 AA; 108395 MW; 2C0EF7FC41CCD060 CRC64;
CC
CC Query Match 29.7%; Score 51; DB 1; Length 962;
CC Best Local Similarity 39.5%; Pred. No. 14;
CC Matches 15; Conservative 0; Mismatches 5; Indels 18; Gaps 2;
CC
CC QY 11 RTPDINAWYAG-----RGIRPVGRP 31
CC Db 703 RTPRSNPATYTGVTTHRELFATPEARKRGYRP-GRF 739
CC
CC RESULT 10
CC Y587_PASMU
CC ID Y587_PASMU STANDARD; PRT; 288 AA.
CC AC Q9CN56;
CC DT 16-OCT-2001 (Rel. 40, Created)
CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC DE Hypothetical protein PM0587.
CC GN PM0587.
CC OS Pasteurella multocida.
CC OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
CC OC Pasteurellaceae; Pasteurella.
CC OX NCBI_TaxID=747;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=Pm70;
CC RX MEDLINE=21145866; PubMed=11248100;
CC RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
CC RT "Complete genomic sequence of Pasteurella multocida Pm70.";
CC Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -1- SIMILARITY: BELONGS TO THE FRUCTOSAMINE KINASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AR006094; AA02671.1; -.  
DR InterPro: IPR005581; Fructosamin\_kin.  
DR Pfam: Pf03881; Fructosamin\_kin; 1.  
DR Hypochemical protein; Transferase; Kinase; Complete proteome.  
SQ SEQUENCE 288 AA; 33778 MW; F4D3F6C26014D940 CRC64;  
KW

```

Query Match      29.1%; Score 50; DB 1; Length 288;
Best Local Similarity 37.5%; Pred.No. 5.4;
Matches 9; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

Qy      5 QHSMEIRTPDINPAWYAGRGIRPV 28
          : : : : :
Db      20 KHKEKHTGEMHEAWIIDGIDQPV 43
          : : : : :

```

```

RESULT 11
UVR_A_RHILO
ID      UVR_A_RHILO      STANDARD;      PRT;      973 AA.
AC      Q98M36;
DT      28-FEB-2003 (Rel. 41, Created)
DT      28-FEB-2003 (Rel. 41, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      UvrABC system protein A (UvrA protein) (Excinuclease ABC subunit A).
GN      UVR_A OR MLR0750.
OS      Rhizobium loti (Mesorhizobium loti).
OC      Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC      Phyllobacteriaceae; Mesorhizobium.
OX      NCBI_TaxID=381;
[1]
RN      SEQUENCE FROM N.A.
RP      STRAIN=MAFF303099;
RX      MEDLINE=21082930; PubMed=11214958;
RA      Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA      Watanabe A., Iidesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA      Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA      Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA      Takeuchi C., Yamada M., Tabata S.;
RT      "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT      Mesorhizobium loti."
RL      DNA Res. 7:331-338(2000).
CC      -1- FUNCTION: The UvrABC repair system catalyzes the recognition and
CC      processing of DNA lesions. UvrA is an ATPase and a DNA-binding
CC      protein. A damage recognition complex composed of 2 uvrA and 2
CC      uvrB subunits scans DNA for abnormalities. When the presence of a
CC      lesion has been verified by uvrB, the uvrA molecules dissociate
CC      (By similarity).
CC      -1- SUBUNIT: Forms a heterotetramer with uvrB during the search for
CC      lesions (By similarity).
CC      -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC      -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. UVR_A SUBFAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/
or send an email to license@isb-sib.ch).

```

```

DR ENBL; AP002995; BAB48277.1; -.
DR HMAP; MF 00205; 1
DR InterPro; IPR003143; ABC transporter.
DR InterPro; IPR004602; UvrA.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 1.
DR TIGRFAMS; TIGR00630; uvra; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 1.

```

|    |                                                                      |
|----|----------------------------------------------------------------------|
| KW | SOS response; Excision nuclease; DNA repair; DNA recombination;      |
| KW | DNA excision; ATP-binding; DNA-binding; Repeat; Zinc; Metal-binding; |
| KW | Zinc-finger; Complete proteome.                                      |
| FT | NP_BIND 34 41 ATP (POTENTIAL) .                                      |
| FT | NP_BIND 662 669 ATP (POTENTIAL) .                                    |
| FT | ZN_FING 761 787 C4-TYPE.                                             |
| SQ | SEQUENCE 973 AA; 107358 MW; 767D7D2C8220057 CRC64;                   |

```

RESULT 12
VUVA RHIME
ID VUVA RHIME STANDARD; PRT; 973 AA.
AC P56899;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DB VUvABC system protein A (VvA protein) (Excinuclease ABC subunit A).
DS VUVA OR R01557 OR SMC01235.
OS Rhizobium meliloti (Sinorhizobium meliloti).
GC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_Taxid=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RC MEDLINE=21396507; PubMed=11491430;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaue V., Masuy D.,
RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,
RA Renard C., Thebault P., Vandenbol M., Weidner S., Galbert F.,
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021."
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).

```

```

RP SEQUENCE OF 1-140 FROM N.A.
RC STRAIN=2021;
RX MEDLINE=99430868; PubMed=10503543;
RC Tapias A., Barbe J.;
RT "Regulation of divergent transcription from the uvrA-sib promoters 1
RT Sinochrobium meliloti";
RL Mol. Gen. Genet. 262:121-130(1999).
CC -1- FUNCTION: The UVrABC repair system catalyzes the recognition and
CC processing of DNA lesions. UvrA is an ATPase and a DNA-binding
CC protein. A damage recognition complex composed of 2 uvrA and 2
CC uvrB subunits scans DNA for abnormalities. When the presence of
CC a lesion has been verified by uvrB, the uvrA molecules dissociate
CC (By similarity).
CC -1- SUBUNIT: Forms a heterotetramer with uvrB during the search for
CC lesions (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. UvrA SUBFAMIL
CC
CC This SWISS-PROT entry is copyright. It is produced through a collabora
CC between the Swiss Institute of Bioinformatics and the EMBL outstac
CC the European Bioinformatics Institute. There are no restrictions o
CC use by non-profit institutions as long as its content is in comm
CC modified and this statement is not removed. Usage by and for comm
CC entities requires a license agreement (See http://www.isb-sib.ch/ann
CC or send an email to license@sib-sib.ch).
CC
CC ENBL; AL591787; CAC461136.1; --
DR ENBL; AF125162; AAF03210.1; --
DR HAMAP; MF_00205; --; 1.
DR InterPro; IPR003439; ABC transporter.

```

DR EMBL; AL591787; CAC46136.1; -;  
DR EMBL; AF125162; AAF03210.1; -;  
DR HAWAP; MF\_00205; -; 1.  
DR InterPro; IPR003439; ABC transporter

DR InterPro; IPR004602; UVRA.  
 DR Pfam; PF00005; ABC\_tran; 2.  
 DR ProDom; PD000006; ABC\_transporter; 1.  
 DR TIGRFAMs; TIGR00630; uvra; 1.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER\_1; 2.  
 DR PROSITE; PS00993; ABC\_TRANSPORTER\_2; 1.  
 KW SOS response; Excision nuclease; DNA repair; DNA recombination;  
 KW DNA excision; ATP-binding; DNA-binding; Repeat; Zinc; Metal-binding;  
 KW Zinc-finger; Complete proteome.  
 FT NP\_BIND 34 41 ATP (POTENTIAL).  
 FT NP\_BIND 662 669 ATP (POTENTIAL).  
 FT ZN\_FING 761 787 C4-TYPE.  
 FT ZN\_FING 761 787 C4-TYPE.  
 FT CONFLICT 19 19 G -> A (IN REF. 2).  
 FT CONFLICT 67 67 F -> S (IN REF. 2).  
 SQ SEQUENCE 973 AA; 107191 MW; 3E1A8B14527A47FE CRC64;  
 Query Match 29.1%; Score 50; DB 1; Length 973;  
 Best Local Similarity 39.5%; Pred. No. 19;  
 Matches 15; Conservative 2; Mismatches 3; Indels 18; Gaps 3;  
 QY 11 RTFDPNPA-----WYAG-----RGIRPVGRF 31  
 DB 716 RTFRSNPATYTGFTPIRDFAGLPEAKRGYQP-GRF 752  
 RESULT 13  
 6PGL\_SYNY3  
 ID 6PGL\_SYNY3 STANDARD; PRT; 239 AA.  
 AC P74618;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE 6-phosphogluconolactonase (EC 3.1.1.31) (6PGL).  
 GN PGL OR DEVB OR SLH1479.  
 OS Synechocystis sp. (strain PCC 6803).  
 OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.  
 OX NCBI\_TaxID=1148;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97061201; PubMed=8905231;  
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,  
 RA Miyajima N., Hirose M., Sugita M., Saito M., Kimura T.,  
 RA Hoshino T., Matsuoka A., Muraki A., Nakazaki N., Naruo K.,  
 RA Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A.,  
 RA Yamada M., Yasuda M., Tabata S.;  
 RT "Sequence analysis of the genome of the unicellular cyanobacterium  
 RT Synechocystis sp. strain PCC6803. II. Sequence determination of the  
 RT entire genome and assignment of potential protein-coding regions.";  
 RL DNA Res. 3:109-136(1996).  
 RN [2]  
 RP SEQUENCE OF 1-19.  
 RX MEDLINE=97443974; PubMed=9298645;  
 RA Sazuka T., Ohara O.;  
 RT "Towards a proteome project of cyanobacterium Synechocystis sp.  
 RT strain PCC6803: linking 130 protein spots with their respective  
 RT genes.";  
 RL Electrophoresis 18:1252-1258(1997).  
 CC -!- FUNCTION: HYDROLYSIS OF 6-PHOSPHOGLUCONOLACTONE TO 6-  
 CC PHOSPHOGLUCONATE.  
 CC -!- CATALYTIC ACTIVITY: 6-phospho-D-glucono-1,5-lactone + H(2)O = 6-  
 CC phospho-D-gluconate.  
 CC -!- PATHWAY: Pentose phosphate pathway; second step.  
 CC -!- SIMILARITY: BELONGS TO THE GLUCOSAMINE/GALACTOSAMINE-6-PHOSPHATE  
 CC ISOMERASE FAMILY. 6-PHOSPHOGLUCONOLACTONASE SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; D90916; BAA18726.1; ALT INIT.  
 DR InterPro; IPR006148; Gluc\_gal\_isom.  
 DR InterPro; IPR005900; Phosphogluconlac.  
 DR Pfam; PF01182; Glucosamine\_iso; 1.  
 DR TIGRFAMs; TIGR01198; pgl; 1.  
 KW Hydrolase; Complete proteome.  
 FT INIT\_MET 0  
 SQ SEQUENCE 239 AA; 26351 MW; 9C64A0A342325917 CRC64;  
 Query Match 28.2%; Score 48.5; DB 1; Length 239;  
 Best Local Similarity 42.3%; Pred. No. 7.5;  
 Matches 11; Conservative 5; Mismatches 9; Indels 1; Gaps 1;  
 QY 5 QHSM-EIRTPDINPAWYAGRGIRPVG 29  
 DB 198 QHLSGIFAPADPQQYPARFIQFG 223  
 RESULT 14  
 Y468\_MYCPN  
 ID Y468\_MYCPN STANDARD; PRT; 1882 AA.  
 AC P75109; Q50317;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Protein MG468 homolog (K05\_orf1882).  
 GN MPN684 OR MP158.  
 OS Mycoplasma pneumoniae.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxID=2104;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 29342 / M129;  
 RX MEDLINE=97105885; PubMed=8948633;  
 RA Himmelreich R., Hilbert H., Piagens H., Pirkl E., Li B.-C.,  
 RA Herrmann R.;  
 RT "Complete sequence analysis of the genome of the bacterium  
 RT Mycoplasma pneumoniae.";  
 RL Nucleic Acids Res. 24:4420-4449(1996).  
 RN [2]  
 RP SEQUENCE OF 1-1848 FROM N.A.  
 RC STRAIN=ATCC 29342 / M129;  
 RX MEDLINE=96177562; PubMed=8604303;  
 RA Hilbert H., Himmelreich R., Piagens H., Herrmann R.;  
 RT "Sequence analysis of 56 kb from the genome of the bacterium  
 RT Mycoplasma pneumoniae comprising the dnaA region, the atp operon and a  
 RT cluster of ribosomal protein genes.";  
 RL Nucleic Acids Res. 24:628-639(1996).  
 RN [3]  
 RP IDENTIFICATION BY MASS SPECTROMETRY.  
 RC STRAIN=ATCC 29342 / M129;  
 RX MEDLINE=21088919; PubMed=11271496;  
 RA Regula J.T., Ueberle B., Boguth G., Goerg A., Schnoelzer M.,  
 RA Herrmann R., Frank R.;  
 RT "Towards a two-dimensional proteome map of Mycoplasma pneumoniae.";  
 RL Electrophoresis 21:3765-3780(2000).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -!- SIMILARITY: SOME, TO MG064.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF000017; AAB95806.1; -;  
 CC EMBL; U34816; AAC43650.1; -;  
 CC PIR; S73484; S73484.  
 CC InterPro; IPR003838; DUF214.  
 CC Pfam; PF02687; FteX; 1.  
 CC Transmembrane; Complete proteome.  
 KW

Search completed: December 3, 2003, 19:12:02  
Job time : 7.25 sec

```
FT TRANSMEM 16 36 POTENTIAL.
FT TRANSMEM 987 1007 POTENTIAL.
FT TRANSMEM 1037 1057 POTENTIAL.
FT TRANSMEM 1080 1100 POTENTIAL.
FT TRANSMEM 1154 1174 POTENTIAL.
FT TRANSMEM 1759 1779 POTENTIAL.
FT TRANSMEM 1807 1827 POTENTIAL.
FT TRANSMEM 1828 1848 POTENTIAL.
FT TRANSMEM 1851 1871 POTENTIAL.
SQ SEQUENCE 1882 AA; 209442 MW; 03CFA4D99A7120ED CRC64;

Query Match 28.2%; Score 48.5; DB 1; Length 1882;
Best Local Similarity 32.3%; Pred. No. 64;
Matches 10; Conservative 6; Mismatches 12; Indels 3; Gaps 1;

QY 1 SRAHQSHMEIRTPDINPAWYAGRGIRPVGRF 31
Db 1188 NRNFNYRLNLQTPTQSGWYA---IQPYSRF 1215
      :| : : :| :| :| :| :| :|
      :| : : :| :| :| :| :| :|

RESULT 15
SAMB_SALTY STANDARD; PRT; 424 AA.
AC P23832;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Samb protein.
GN SAMB.
OS Salmonella typhimurium.
OG plasmid 60-mDa cryptic.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2;
RX MEDLINE=91123176; PubMed=1991707;
RA Nohmi T., Hakura A., Nakai Y., Watanabe M., Murayama S.Y.,
RA Sofuni T.;
RT "Salmonella typhimurium has two homologous but different umuDC
RT operons: cloning of a new umuDC-like operon (samAB) present in a
RT 60-megadalton cryptic plasmid of S. typhimurium.";
RL J. Bacteriol. 173:1051-1053(1991).
CC -1- FUNCTION: INVOLVED IN UV PROTECTION AND MUTATION.
CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-Y FAMILY.
CC -1- SIMILARITY: Contains 1 umuc domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D90202; BAA14226.1; -.
DR PIR; B38176; B38176.
DR HAMAP; MF 01113; atypical; 1.
DR InterPro; IPR001126; UMUC_like.
DR Pfam; PF00817; IMS; 1.
DR PROSITE; PS50173; UMUC; 1.
KW Plasmid; SOS mutagenesis; DNA repair.
FT DOMAIN 2 189 UMUC.
SQ SEQUENCE 424 AA; 47727 MW; FF8C47476CC58A2B CRC64;

Query Match 27.9%; Score 48; DB 1; Length 424;
Best Local Similarity 50.0%; Pred. No. 16;
Matches 9; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 10 IRTPDINPAWYAGRGIRP 27
Db 381 INHPGKGKIWFAGRGIAIP 398
      | :| :| :| :| :| :| :| :| :|
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 19:03:25 ; Search time 26.5 Seconds  
(without alignments)  
301.873 Million cell updates/sec

Title: US-09-868-885b-3

Perfect score: 172

Sequence: 1 SRAHQSHMEIRTPDINPAWYAGRGIRPVGRF 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL 23:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID  | Description         |
|------------|-------|-------------|--------|--------|---------------------|
| 1          | 172   | 100.0       | 98     | Q8WN12 | Q8wn12 ovis aries   |
| 2          | 104   | 60.5        | 117    | Q9W624 | Q9w624 carassius a  |
| 3          | 62    | 36.0        | 692    | Q9L8J6 | Q9l8j6 rhodospirill |
| 4          | 57    | 33.1        | 54     | Q9UJF9 | Q9ujf9 homo sapien  |
| 5          | 57    | 33.1        | 355    | Q95YJ8 | Q95yj8 ciona savig  |
| 6          | 57    | 33.1        | 465    | Q60687 | Q60687 homo sapien  |
| 7          | 57    | 33.1        | 465    | Q8WW85 | Q8ww85 homo sapien  |
| 8          | 56.5  | 32.8        | 303    | Q8YWC7 | Q8ywc7 anabaena sp  |
| 9          | 55    | 32.0        | 1236   | Q9JPA4 | Q9jpa4 rhodocyclu   |
| 10         | 54.5  | 31.7        | 503    | Q9KRY1 | Q9kry1 vibrio chol  |
| 11         | 53.5  | 31.1        | 176    | Q08689 | Q08689 saccharomyc  |
| 12         | 53    | 30.8        | 376    | Q9RYP2 | Q9ryp2 deinococcus  |
| 13         | 53    | 30.8        | 467    | Q8R054 | Q8r054 mus musculu  |
| 14         | 53    | 30.8        | 468    | Q8K1F8 | Q8k1f8 mus musculu  |
| 15         | 52    | 30.2        | 73     | Q8U515 | Q8u515 agrobacteri  |
| 16         | 52    | 30.2        | 308    | Q8UC89 | Q8uc89 agrobacteri  |

|    |      |      |      |    |        |                    |
|----|------|------|------|----|--------|--------------------|
| 17 | 52   | 30.2 | 419  | 16 | Q8UKP2 | Q8ukp2 agrobacteri |
| 18 | 52   | 30.2 | 440  | 11 | Q8CBT2 | Q8cbt2 mus musculu |
| 19 | 52   | 30.2 | 447  | 16 | Q8FOP1 | Q8fop1 corynebacte |
| 20 | 52   | 30.2 | 848  | 11 | Q8C720 | Q8c720 mus musculu |
| 21 | 52   | 30.2 | 1253 | 2  | Q8RTV7 | Q8rtv7 uncultured  |
| 22 | 52   | 30.2 | 1292 | 2  | Q9ZGES | Q9zges heliobacill |
| 23 | 52   | 30.2 | 3567 | 11 | Q9ES77 | Q9es77 mus musculu |
| 24 | 51.5 | 29.9 | 327  | 10 | Q94KU9 | Q94ku9 brassica ca |
| 25 | 51.5 | 29.9 | 327  | 10 | Q94PZ9 | Q94pz9 brassica ca |
| 26 | 51.5 | 29.9 | 664  | 16 | Q91LW4 | Q91lw4 pseudomonas |
| 27 | 51   | 29.7 | 184  | 16 | Q8PER0 | Q8per0 xanthomonas |
| 28 | 51   | 29.7 | 294  | 5  | Q19530 | Q19530 caenorhabdi |
| 29 | 51   | 29.7 | 315  | 2  | Q87474 | Q87474 burkholderi |
| 30 | 51   | 29.7 | 333  | 4  | Q96SD4 | Q96sd4 homo sapien |
| 31 | 51   | 29.7 | 367  | 4  | Q8N6Q2 | Q8n6q2 homo sapien |
| 32 | 51   | 29.7 | 367  | 4  | Q8IXT2 | Q8ixt2 homo sapien |
| 33 | 51   | 29.7 | 370  | 11 | Q8CGW9 | Q8cgw9 mus musculu |
| 34 | 51   | 29.7 | 430  | 17 | Q27142 | Q27142 methanobact |
| 35 | 51   | 29.7 | 514  | 5  | Q9VRV3 | Q9vrv3 drosophila  |
| 36 | 51   | 29.7 | 689  | 11 | Q8R3R1 | Q8r3r1 mus musculu |
| 37 | 51   | 29.7 | 767  | 11 | Q8JZN2 | Q8jzn2 mus musculu |
| 38 | 51   | 29.7 | 798  | 11 | Q91W50 | Q91w50 mus musculu |
| 39 | 51   | 29.7 | 1563 | 5  | Q8T4U2 | Q8t4u2 manduca sex |
| 40 | 51   | 29.7 | 1845 | 17 | Q8TT57 | Q8tt57 methanosarc |
| 41 | 50.5 | 29.4 | 292  | 10 | Q8S7E3 | Q8s7e3 oryza sativ |
| 42 | 50.5 | 29.4 | 374  | 16 | Q8ZLE1 | Q8zle1 salmonella  |
| 43 | 50.5 | 29.4 | 420  | 2  | Q93LZ7 | Q93lz7 streptomyce |
| 44 | 50.5 | 29.4 | 790  | 10 | Q9M371 | Q9m371 arabidopsis |
| 45 | 50   | 29.1 | 217  | 16 | Q8PQ42 | Q8pq42 xanthomonas |

#### ALIGNMENTS

#### RESULT 1

Q8WN12 PRELIMINARY; PRT; 98 AA.

AC Q8WN12;  
DT 01-MAR-2002 (TRENBLrel. 20, Created)  
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)  
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)  
DE Preprolactin-releasing peptide.  
OS Ovis aries (Sheep).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Caprinae; Ovis.  
OX NCBI\_TaxID=9940;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Curlewis J.D.; Kusters D.H.L.; Barclay J.L.; Anderson S.T.;  
RT "Prolactin-releasing peptide (PrRP) in the ewe: cDNA cloning, mRNA  
RT distribution and effects on prolactin secretion in vitro and in  
RT vivo."  
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF450453; AAL47178.1;  
SQ SEQUENCE 98 AA; 10513 MW; 2A53331ED62CAAB5 CRC64;

Query Match 100.0%; Score 172; DB 6; Length 98;  
Best Local Similarity 100.0%; Pred. No. 9.2e-17;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQSHMEIRTPDINPAWYAGRGIRPVGRF 31  
|||||  
DB 23 SRAHQSHMEIRTPDINPAWYAGRGIRPVGRF 53  
|||||

#### RESULT 2

Q9W624 PRELIMINARY; PRT; 117 AA.

AC Q9W624;  
DT 01-NOV-1999 (TRENBLrel. 12, Created)  
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)  
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)

```

RN      [1]
RP      SEQUENCE FROM N.A.
RA      Laviot S.;
RL      Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AL035608; CAB55682.1; -
DL      NON TER      54
FT      SEQUENCE      54 AA; 6110 MW; E2P3C39F7B961A9P CRC64;
SQ

Query Match      33.1%; Score 57; DB 4; Length 54;
Best Local Similarity 56.2%; Pred. NO. 0.98;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY      12 TPDINPAWYAGRGIRP 27
      ||| : |||||
DB      18 TPAVTPWYAGSGYYP 33

RESULT 5
Q95YJ8 PRELIMINARY; PRT; 355 AA.
ID
AC Q95YJ8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE zic related protein 1a.
GN CS-ZICRIA.
OS Clona savigny1.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
OC Phlebobranchia; Clonidae; Clona.
OX NCBI_TaxID=51511;
RN
RP      SEQUENCE FROM N.A.
RA      Inai K.S., Satoh N., Satou Y.;
RL      "Clona savigny1 genes.";
RL      Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AB057747; BAB68356.1; -
DR      InterPro; IPR007087; Znf_C2H2.
DR      Pfam; PF00096; zf-C2H2; 4.
DR      SMART; SM00355; Znf_C2H2; 4.
DR      PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
DR      PROSITE; PS0157; ZINC_FINGER_C2H2_2; 4.
KW Metal-binding; Zinc; Zinc-finger.
SQ      SEQUENCE 355 AA; 40876 MW; E58F5DEDD812E8AC CRC64;

Query Match      33.1%; Score 57; DB 5; Length 355;
Best Local Similarity 40.0%; Pred. No. 7;
Matches 12; Conservative 7; Mismatches 7; Indels 4; Gaps 1;

QY      2 RAHQHSEWIRTPDIN----PAWYAGRGIRP 27
      :|||::: :| :|||::|
DB      27 QAHQHSIDSKPQLNSVPSAAYAGYGMP 56

RESULT 6
O60687 PRELIMINARY; PRT; 465 AA.
ID
AC O60687;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Sushi-repeat protein (Sushi-repeat containing protein).
GN SRPUL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN
RP      SEQUENCE FROM N.A.
RA      Kuroshaw H., Inukai T., Inaba T., Gai K., Chang K.-S., Sinjo T.,
RL      Rakesraw K.M., Naeve C.W., Look T.A.;
RL      Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
[2]
RN      SEQUENCE FROM N.A.
RP

```

RA Huang C.-H., Chen H., Peng J., Chen Y.;  
 RT "Cloning and characterization of the sushi-repeat containing protein  
 RT (SRP) as a novel interaction partner of Rh type C glycoprotein  
 RT (RhCG).";  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF060567; AAC15765.1; -.  
 DR EMBL; AF393649; AAM73693.1; -.  
 DR InterPro; IPR001128; Cytochrome\_P450.  
 DR InterPro; IPR003410; Hyalin.  
 DR InterPro; IPR000436; Sushi\_SCR\_CCP.  
 DR Pfam; PF02494; HYR; 1.  
 DR Pfam; PF00084; sushi; 3.  
 DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
 SQ SEQUENCE 465 AA; 52971 MW; 4D752B187FF3EFB8 CRC64;

Query Match 33.1%; Score 57; DB 4; Length 465;

Best Local Similarity 56.2%; Pred. No. 9.2; Mismatches 6; Indels 0; Gaps 0;

QY 12 TPDINPAWYAGRGIRP 27  
 || : |||||  
 Db 18 TPAVFTWYAGSGYYP 33

## RESULT 7

Q8W85 PRELIMINARY; PRT; 465 AA.  
 ID Q8W85;  
 AC Q8W85;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Sushi-repeat protein.  
 OS Homo sapiens (Human)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placenta;  
 RA Strausberg R.;  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC020733; AAH20733.1; -.  
 DR InterPro; IPR001128; Cytochrome\_P450.  
 DR InterPro; IPR003410; Hyalin.  
 DR InterPro; IPR000436; Sushi\_SCR\_CCP.  
 DR Pfam; PF02494; HYR; 1.  
 DR Pfam; PF00084; sushi; 3.  
 DR SMART; SM00032; CCP; 3.  
 DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
 SQ SEQUENCE 465 AA; 52957 MW; 3D7229487DA1B8BD CRC64;

Query Match 33.1%; Score 57; DB 4; Length 465;

Best Local Similarity 56.2%; Pred. No. 9.2; Mismatches 6; Indels 0; Gaps 0;

QY 12 TPDINPAWYAGRGIRP 27  
 || : |||||  
 Db 18 TPAVFTWYAGSGYYP 33

## RESULT 8

Q8W85 PRELIMINARY; PRT; 303 AA.  
 ID Q8W85;  
 AC Q8W85;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Hypothetical protein A11686.  
 GN A11686.  
 OS Anabaena sp. (strain PCC 7120).  
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.  
 OX NCBI\_TaxID=103690;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=21595285; PubMed=11759840;  
 RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,  
 RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,  
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,  
 RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,  
 RA Yasuda M., Tabata S.;  
 RT "Complete genomic sequence of the filamentous nitrogen-fixing  
 RT cyanobacterium Anabaena sp. strain PCC 7120.";  
 RL DNA Res. 8:205-213(2001).  
 DR EMBL; AF003586; BAB78052.1; -.  
 DR InterPro; IPR004843; M-ppetrase.  
 DR Pfam; PF00149; Metallophos; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 303 AA; 34449 MW; 685700B2127BE987 CRC64;

Query Match 32.8%; Score 56.5; DB 16; Length 303;

Best Local Similarity 51.6%; Pred. No. 7; Mismatches 8; Indels 5; Gaps 2;

QY 4 HOHSMETRTPDINPAWY----AGRGIRPVGR 30  
 || : |||||  
 Db 226 HEHSYE-RTAIDGTTLTCGAGAGNRPVGR 255

## RESULT 9

Q8JPA4 PRELIMINARY; PRT; 1236 AA.  
 ID Q8JPA4;  
 AC Q8JPA4;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Mg protoporphyrin methyl transferase.  
 GN ECHH.  
 OS Rhodocyclus gelatinosus (Rhodopseudomonas gelatinosa).  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
 OC Comamonadaceae; Rubrivivax.  
 OX NCBI\_TaxID=28068;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=IL144;  
 RA Nagashima K.V., Shimada K., Matsuura K.;  
 RT "Phylogenetic analysis of photosynthetic genes of Rhodocyclus  
 RT gelatinosus: Possibility of horizontal gene transfer in purple  
 RT bacteria.";  
 RL Photosyn. Res. 36:185-191(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=IL144;  
 RX MEDLINE=94132007; PubMed=8300574;  
 RA Nagashima K.V., Matsuura K., Ohyama S., Shimada K.;  
 RT "Primary structure and transcription of genes encoding B870 and  
 RT photosynthetic reaction center apoproteins from Rubrivivax  
 RT gelatinosus.";  
 RL J. Biol. Chem. 269:2477-2484(1994).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=IL144;  
 RX MEDLINE=20031519; PubMed=10563807;  
 RA Menin L., Yoshida M., Jaquinod M., Nagashima K.V., Matsuura K.,  
 RA Parot P., Vermiglio A.;  
 RT "Dark aerobic growth conditions induce the synthesis of a high

RP SEQUENCE FROM N.A.  
 RX MEDLINE=20031519; PubMed=10563807;  
 RA Menin L., Yoshida M., Jaquinod M., Nagashima K.V., Matsuura K.,  
 RA Parot P., Vermiglio A.;  
 RT "Dark aerobic growth conditions induce the synthesis of a high

RT midpoint potential cytochrome c8 in the photosynthetic bacterium  
 RL Rubrivivax gelatinosus.";  
 DR Biochemistry 38:15238-15244 (1999).  
 DR EMBL; AB034704; BAA94057.1; -.  
 DR InterPro; IPR003672; Cbn/Mg\_chitase.  
 DR Pfam; PF02514; cbn-Mg\_chel; 1.  
 KW Transferase.  
 SQ SEQUENCE 1236 AA; 134729 MW; 84051C045638520C CRC64;

Query Match 32.0%; Score 55; DB 2; Length 1236;  
 Best Local Similarity 34.4%; Pred. No. 49;  
 Matches 11; Conservative 6; Mismatches 9; Indels 6; Gaps 1;

Qy 3 AHQSHMEIRTPDINPAWAG-----RGIRPV 28  
 Db 1112 SEQVALETRMLNPKYEGMLEHGYEGVRQI 1143

RESULT 10  
 Q9KRY1 PRELIMINARY; PRT; 503 AA.  
 AC Q9KRY1;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Sun/nucleolar protein family protein.  
 GN VC1502.  
 OS Vibrio cholerae.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
 OC Vibrionaceae; Vibrio.  
 OX NCBI\_TaxID=666;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RP STRAIN=E1 Tor N16961 / Serotype O1;  
 RX MEDLINE=20406833; PubMed=10952301;  
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,  
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,  
 RA Gill S.R., Nelson K.S., Read T.D., Tettelin H., Richardson D.,  
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,  
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,  
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
 RA Fraser C.M.;  
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio  
 RT cholerae.";  
 RL Nature 406:477-483 (2000).  
 DR EMBL; AE004228; AAF94657.1; -.  
 DR TIGR; VC1502; -.  
 DR InterPro; IPR00051; SAM bind.  
 DR Pfam; PF01189; Noll\_Nop2\_Sun; 1.  
 DR TIGRFAMs; TIGR00446; nop2p; 1.  
 DR PROSITE; PS01153; NOLL\_NOP2\_SUN; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 503 AA; 55997 MW; 2ABD94A2356C9E48 CRC64;

Query Match 31.7%; Score 54.5; DB 16; Length 503;  
 Best Local Similarity 36.7%; Pred. No. 23;  
 Matches 11; Conservative 6; Mismatches 12; Indels 1; Gaps 1;

Qy 1 SRAHQSHMEIRTPDINPAWAGRGIRPVGR 30  
 Db 435 SSSASHSVELDTQAR-EWFGMDVRPEQG 463

RESULT 11  
 Q08689 PRELIMINARY; PRT; 176 AA.  
 AC Q08689;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Chromosome XV reading frame ORF YOR253W.  
 GN ARD2 OR YOR253W.

OS Saccharomyces cerevisiae (Baker's Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RP MIPS;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RC SEQUENCE FROM N.A.  
 RP MEDLINE=97298311; PubMed=9153759;  
 RX Jaumiaux J.C., Poiry R.;  
 RT "Sequencing analysis of a 36.8 kb fragment of yeast chromosome XV  
 RT reveals 26 open reading frames including SEC63, CDC31, SUG2, GCD1,  
 RT RBL2, PNT1, PAC1 and VPH1.";  
 RT Yeast 13:483-487 (1997).  
 DR EMBL; Z75161; CAA39475.1; -.  
 DR SGD; S0005779; ARD2.  
 DR InterPro; IPR000345; CytC heme bind.  
 DR InterPro; IPR000182; GCNSAcetyltransf.  
 DR Pfam; PF00583; Acetyltransf; 1.  
 DR PROSITE; PS00190; CYTOCHROME C; 1.  
 SQ SEQUENCE 176 AA; 19727 MW; 4F09DC597A690BA0 CRC64;

Query Match 31.1%; Score 53.5; DB 3; Length 176;  
 Best Local Similarity 31.2%; Pred. No. 11;  
 Matches 10; Conservative 5; Mismatches 14; Indels 3; Gaps 1;

Qy 1 SRAHQSHMEIRTP---DINPAWAGRGIRPVG 29  
 Db 120 SECHQNVFVYLPVAVDDLTQKQFIAHGFEQVG 151

RESULT 12  
 Q9RYP2 PRELIMINARY; PRT; 376 AA.  
 AC Q9RYP2;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Adenine deaminase-related protein.  
 GN DRA0268.  
 OS Deinococcus radiodurans.  
 OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;  
 OC Deinococcaceae; Deinococcus.  
 OX NCBI\_TaxID=1299;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RP STRAIN=R1;  
 RX MEDLINE=20036896; PubMed=10567266;  
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,  
 RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,  
 RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,  
 RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,  
 RA Makarov K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,  
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,  
 RA Fraser C.M.;  
 RT "Genome sequence of the radioresistant bacterium Deinococcus  
 RT radiodurans R1.";  
 RL Science 286:1571-1577 (1999).  
 DR EMBL; AE001863; AAF12376.1; -.  
 DR TIGR; DRA0268; -.  
 KW Complete proteome.  
 SQ SEQUENCE 376 AA; 39845 MW; 7AB7FF32F8C45651 CRC64;

Query Match 30.8%; Score 53; DB 16; Length 376;  
 Best Local Similarity 40.0%; Pred. No. 27;  
 Matches 12; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

Qy 1 SRAHQSHMEIRTPDINPAWAGRGIRPVGR 30  
 Db 177 ARALHAPDVSRPDHRACGAGQRRRPAQR 206



```

RESULT 13
Q8R054
ID Q8R054 PRELIMINARY; PRT; 467 AA.
AC Q8R054;
DT 01-JUN-2002 (TRENBLrel. 21, Created)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Similar to sushi-repeat protein.
GN 1110039C07RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC028307; AAH28307.1;
DR MGD; MGI:1916042; 1110039C07RIK.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR003410; Hyalin.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF02494; HYR; 1.
DR Pfam; PF00084; sushi; 3.
DR SMART; SM00032; CCP; 3.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
DR SEQUENCE 467 AA; 53009 MW; BB4C01C7B6118BE0 CRC64;

Query Match 30.8%; Score 53; DB 11; Length 467;
Best Local Similarity 53.3%; Pred. No. 34;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 13 PDINPAWYAGRGIRP 27
DB 19 PAVTPTWYAGSGYSP 33

RESULT 14
Q8K1F8
ID Q8K1F8 PRELIMINARY; PRT; 468 AA.
AC Q8K1F8;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Sushi-repeat containing protein.
GN 1110039C07RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Huang C.-H., Chen H., Peng J., Chen Y.;
RT "Cloning and characterization of the sushi-repeat containing protein
RT (SRP) as a novel interaction partner of Rh type C glycoprotein
RT (RhCG).";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF393640; AAM73691.1;
DR MGD; MGI:1916042; 1110039C07RIK.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR003410; Hyalin.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF02494; HYR; 1.
DR Pfam; PF00084; sushi; 3.
DR SMART; SM00032; CCP; 3.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
DR SEQUENCE 468 AA; 53180 MW; 151A952070D040D4 CRC64;

Query Match 30.8%; Score 53; DB 11; Length 468;
Best Local Similarity 53.3%; Pred. No. 34;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

```

```

QY 13 PDINPAWYAGRGIRP 27
DB 20 PAVTPTWYAGSGYSP 34

RESULT 15
Q8U515
ID Q8U515 PRELIMINARY; PRT; 73 AA.
AC Q8U515;
DT 01-JUN-2002 (TRENBLrel. 21, Created)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE AGR_L109P.
GN AGR_L109.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houniel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Planagan C., Crowell C., Gursen J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58.";
RL Science 294:2323-2328(2001).
DR EMBL; AE008203; AAK88619.1;
DR SEQUENCE 73 AA; 8219 MW; 02A6F70FA651F2AB CRC64;

Query Match 30.2%; Score 52; DB 16; Length 73;
Best Local Similarity 47.6%; Pred. No. 6.9;
Matches 10; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 3 AQHSHMEIRTPDINPAWYAGR 23
DB 23 AHTHRDEVRSACISVRWLAGR 43

Search completed: December 3, 2003, 19:14:01
Job time : 27.5 secs

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 18:57:14 ; Search time 33.75 Seconds  
(without alignments)  
145.793 Million cell updates/sec

Title: US-09-868-885b-18

Perfect score: 174

Sequence: 1 SRAHQSMETRTFDINPAWYTCGRGIRPVGRF 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_19Jun03.\*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description                 |
|------------|-------|-------------|--------|-------|-----------------------------|
| 1          | 174   | 100.0       | 31     | 18    | AAW31384 Rat type G protein |
| 2          | 174   | 100.0       | 31     | 20    | AAW97233 Rat type ligand po |
| 3          | 174   | 100.0       | 31     | 20    | AAW87614 Rat 19p2 ligand.   |
| 4          | 174   | 100.0       | 31     | 20    | AAW95173 Murine pituitary-d |
| 5          | 174   | 100.0       | 31     | 20    | AAW95174 Murine pituitary-d |
| 6          | 174   | 100.0       | 31     | 21    | AAW10355 Rat oxytocin secre |
| 7          | 174   | 100.0       | 31     | 21    | AAW87504 Rat prolactin-rele |
| 8          | 174   | 100.0       | 31     | 21    | AAW49292 19p2 ligand peptid |
| 9          | 174   | 100.0       | 31     | 22    | AAW62524 Rat CRH releasing  |

|    |     |       |    |    |                             |
|----|-----|-------|----|----|-----------------------------|
| 10 | 174 | 100.0 | 31 | 22 | AAW90993 Prolactin relea-   |
| 11 | 174 | 100.0 | 31 | 23 | AAW26400 Rat PRP-31 peptio  |
| 12 | 174 | 100.0 | 31 | 24 | ABU60826 Peptide production |
| 13 | 174 | 100.0 | 31 | 24 | ABU60837 Peptide production |
| 14 | 174 | 100.0 | 32 | 18 | AAW31385 Rat type G protein |
| 15 | 174 | 100.0 | 32 | 21 | AAW10356 Rat oxytocin secre |
| 16 | 174 | 100.0 | 32 | 22 | AAW62525 Rat CRH releasing  |
| 17 | 174 | 100.0 | 32 | 24 | ABU60838 Peptide production |
| 18 | 174 | 100.0 | 33 | 18 | AAW31386 Rat type G protein |
| 19 | 174 | 100.0 | 33 | 21 | AAW10357 Rat oxytocin secre |
| 20 | 174 | 100.0 | 33 | 22 | AAW62526 Rat CRH releasing  |
| 21 | 174 | 100.0 | 33 | 24 | ABU60839 Peptide production |
| 22 | 174 | 100.0 | 82 | 20 | AAW95172 Murine pituitary-d |
| 23 | 174 | 100.0 | 83 | 18 | AAW31383 Rat type G protein |
| 24 | 174 | 100.0 | 83 | 20 | AAW97225 19p2 ligand po     |
| 25 | 174 | 100.0 | 83 | 21 | AAW10354 Rat oxytocin secre |
| 26 | 174 | 100.0 | 83 | 22 | AAW62523 Rat CRH releasing  |
| 27 | 164 | 94.3  | 31 | 22 | AAW73370 bPRP31 peptide, u  |
| 28 | 163 | 93.7  | 31 | 18 | AAW31371 Bovine G protein-c |
| 29 | 163 | 93.7  | 31 | 20 | AAW97218 Bovine pituitary-d |
| 30 | 163 | 93.7  | 31 | 20 | AAW87613 Bovine 19p2 ligand |
| 31 | 163 | 93.7  | 31 | 20 | AAW95188 Bovine pituitary-d |
| 32 | 163 | 93.7  | 31 | 21 | AAW10347 Bovine oxytocin se |
| 33 | 163 | 93.7  | 31 | 21 | AAW49290 19p2 ligand peptid |
| 34 | 163 | 93.7  | 31 | 21 | AAW49298 19p2 ligand peptid |
| 35 | 163 | 93.7  | 31 | 22 | AAW62516 Bovine CRH releas  |
| 36 | 163 | 93.7  | 31 | 23 | AAW26399 Bovine PRP-31 pep  |
| 37 | 163 | 93.7  | 31 | 24 | ABU60825 Peptide production |
| 38 | 163 | 93.7  | 31 | 24 | ABU60831 Peptide production |
| 39 | 163 | 93.7  | 32 | 18 | AAW31372 Bovine G protein-c |
| 40 | 163 | 93.7  | 32 | 20 | AAW95189 Bovine pituitary-d |
| 41 | 163 | 93.7  | 32 | 21 | AAW10348 Bovine oxytocin se |
| 42 | 163 | 93.7  | 32 | 22 | AAW62517 Bovine CRH releas  |
| 43 | 163 | 93.7  | 32 | 24 | ABU60832 Peptide production |
| 44 | 163 | 93.7  | 33 | 18 | AAW31373 Bovine G protein-c |
| 45 | 163 | 93.7  | 33 | 20 | AAW95190 Bovine pituitary-d |

#### ALIGNMENTS

RESULT 1

AAW31384

ID AAW31384 standard; Peptide; 31 AA.

XX AAW31384;

XX 06-APR-1998 (first entry)

XX Rat type G protein-coupled receptor ligand fragment 1.

XX G protein-coupled receptor; ligand binding; pharmaceutical;  
XX modulator; pituitary; central nervous system; pancreas; prophylactic;  
XX therapeutic agent.

OS Rat sp.

XX WQ9724436-A2.

XX 10-JUL-1997.

XX 26-DEC-1996; 96WO-JP03821.

XX 18-SEP-1996; 96JP-0246573.

PR 28-DEC-1995; 95JP-0343371.

PR 15-MAR-1996; 96JP-0059419.

PR 12-AUG-1996; 96JP-0211805.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;

PI Kawamata Y, Kitada C;

XX

DR WPI; 1997-363672/33.  
 DR N-PSDB; AAV02421.  
 XX  
 PT Ligand peptide for G protein-coupled receptor - acts by modulating  
 PT function in the central nervous system, pancreas and pituitary gland  
 XX  
 PS  
 PS Claim 2; Page 179; 258pp; English.  
 XX  
 CC This sequence represents a peptide fragment from a novel rat type  
 CC ligand polypeptide corresponding to amino acid residues 22 to 52 of the  
 CC sequence represented in AAW31383 and is used in an assay to monitor  
 CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
 CC compositions containing this ligand may be used as a pituitary function  
 CC modulator, a central nervous system modulator or a pancreatic function  
 CC modulator. This ligand could have specific applications as a  
 CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
 CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
 CC trauma, growth hormone secretory disease, hyper- and polyphagia,  
 CC hyperlipidaemia, hypercholesterolaemia, hyperglycaemia,  
 CC hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease,  
 CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
 CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
 CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
 CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
 CC oligogalactia. Assays can also be developed to screen compounds which are  
 CC capable of altering the binding activity of the ligand affecting  
 CC activation of the G protein-coupled receptor protein.  
 XX  
 SQ Sequence 31 AA;  
 Query Match 100.0%; Score 174; DB 18; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-18;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 SRAHQHSMETRTDPINPAWYTGIRPVGRF 31  
 Db 1 SRAHQHSMETRTDPINPAWYTGIRPVGRF 31  
 RESULT 2  
 AAW97233  
 ID AAW97233 standard; peptide; 31 AA.  
 XX  
 AC AAW97233;  
 XX  
 DT 06-MAY-1999 (first entry)  
 XX  
 DE Rat type ligand polypeptide fragment.  
 XX  
 KW Rat type ligand; modulation; prolactin secretion;  
 KW G protein-coupled receptor; GPCR; hypovarianism; gonocyst cacogenesis;  
 KW menopausal syndrome; euthyroid; hypometabolism; lactation;  
 KW pituitary adenomatosis; brain tumour; emmenopathy; autoimmune disease;  
 KW prolactinoma; infertility; impotence; amenorrhea; galactorrhea;  
 KW acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;  
 KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;  
 KW contraceptive; placental function; choriocarcinoma; hydatic mole;  
 KW irruption mole; abortion; unthrifty fetus; abnormal saccharometabolism;  
 KW abnormal lipidmetabolism; oxytocia.  
 XX  
 OS Rattus sp.  
 XX  
 PN W09858962-A1.  
 XX  
 PD 30-DEC-1998.  
 XX  
 PP 22-JUN-1998; 98WO-JP02765.  
 XX  
 PR 23-JUN-1997; 97JP-0165437.  
 XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PI Fujii R, Hinuma S, Kawamata Y, Matsumoto H;  
 WPI; 1997-363672/33.  
 N-PSDB; AAV02421.  
 Ligand peptide for G protein-coupled receptor - acts by modulating  
 function in the central nervous system, pancreas and pituitary gland  
 Claim 2; Page 179; 258pp; English.  
 This sequence represents a peptide fragment from a novel rat type  
 ligand polypeptide corresponding to amino acid residues 22 to 52 of the  
 sequence represented in AAW31383 and is used in an assay to monitor  
 ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
 compositions containing this ligand may be used as a pituitary function  
 modulator, a central nervous system modulator or a pancreatic function  
 modulator. This ligand could have specific applications as a  
 prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
 syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
 trauma, growth hormone secretory disease, hyper- and polyphagia,  
 hyperlipidaemia, hypercholesterolaemia, hyperglycaemia,  
 hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease,  
 Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
 transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
 acute myocardial infarction, infertility, spinocerebellar degeneration,  
 bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
 oligogalactia. Assays can also be developed to screen compounds which are  
 capable of altering the binding activity of the ligand affecting  
 activation of the G protein-coupled receptor protein.  
 Sequence 31 AA;  
 Query Match 100.0%; Score 174; DB 18; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-18;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 SRAHQHSMETRTDPINPAWYTGIRPVGRF 31  
 Db 1 SRAHQHSMETRTDPINPAWYTGIRPVGRF 31  
 RESULT 3  
 AAW87614  
 ID AAW87614 standard; Peptide; 31 AA.  
 XX  
 AC AAW87614;  
 XX  
 DT 29-MAR-1999 (first entry)  
 XX  
 DE Rat 19P2 ligand.  
 XX  
 KW 19P2 ligand; G protein coupled receptor; pituitary;  
 KW prolactin releasing peptide; rat; dementia; breast cancer;  
 KW therapy.  
 XX  
 OS Rattus sp.  
 XX  
 PN EP887417-A2.  
 XX  
 PD 30-DEC-1998.  
 XX  
 PP 25-JUN-1998; 98EP-0111725.  
 XX  
 PR 27-JUN-1997; 97JP-0172118.  
 XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PI Moriya T, Nishimura O, Suenaga M, Tanaka Y;  
 WPI; 1999-047884/05.  
 Producing a 19P2 pituitary G protein receptor ligand - by cleavage  
 of a fusion protein, useful for preventing and treating dementia,  
 breast cancer, renal failure and autoimmune disease  
 Claim 5; Page 34; 56pp; English.

XX This is the amino acid sequence of the rat pituitary G  
 CC protein-coupled receptor ligand 192L. A method suitable for  
 CC commercial high-level production of 192L comprises expressing  
 CC the ligand in host cells as a recombinant fusion protein e.g. with  
 CC human basic fibroblast growth factor (see AAV83794-95) that has  
 CC been modified to include an N-terminal cysteine residue. The  
 CC ligand is released from the fusion by cyanylation followed by  
 CC ammonolysis. 192L has prolactin secretion-stimulating and (at  
 CC high doses) prolactin secretion-inhibiting properties. It can be  
 CC used in the treatment and prevention of various diseases including:  
 CC senile dementia, cerebrovascular dementia, and dementia associated  
 CC with: neurological disorders (e.g. Alzheimer's disease, Parkinson's  
 CC disease, Pick's disease, Huntington's disease), infectious diseases  
 CC (e.g. Creutzfeldt-Jakob's), endocrine or metabolic disease or  
 CC intoxicosis (e.g. hypothyroidism, vitamin B12 deficiency, alcoholism,  
 CC diseases (e.g. brain tumour), traumatic diseases (e.g. chronic  
 CC subarachnoid haemorrhage, and other types of dementia, depression,  
 CC hyperactive child syndrome (microencephalopathy) and disturbance of  
 CC consciousness. It is also useful for prevention and treatment of  
 CC diseases associated with prolactin hypo and hypersecretion  
 CC respectively, including: hyperprolactinaemia, pituitary adenoma,  
 CC breast cancer, infertility, impotence and autoimmune disease  
 CC (hypersecretion disorders), and seminal vesicle hypoplasia  
 CC osteoporosis, menopausal syndrome and renal failure (hyposecretion  
 CC disorders). The 192L polypeptide/amide is also useful as a test  
 CC reagent for study of the prolactin secretory function or as a  
 CC lactagogue in mammalian farm animals.

XX Sequence 31 AA;

Query Match 100.0%; Score 174; DB 20; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-18;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQSMETRTPDINPAWYTGIRPVGRF 31  
 |||||  
 Db 1 SRAHQSMETRTPDINPAWYTGIRPVGRF 31

## RESULT 4

AAW95173  
 ID AAW95173 standard; peptide; 31 AA.

XX  
 AC AAW95173;

XX 10-MAR-1999 (first entry)

XX Murine pituitary-derived ligand mature polypeptide sequence.

XX Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;  
 KW GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas;  
 KW tissue; screen; therapeutic; binding; senile dementia; ligand; murine;  
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;  
 KW Creutzfeld-Jakob disease; poisoning; schizophrenia; growth hormone;  
 KW secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;  
 KW gene therapy; transgenic animal.

XX Mus sp.

XX WO9849295-A1.

XX 05-NOV-1998.

XX 27-APR-1998; 98WO-JP01923.

XX 28-APR-1997; 97JP-0109974.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Fukusumi S, Hiruma S;

XX

DR WPI; 1999-009423/01.

XX New polypeptide ligand for orphan G protein coupled receptors - used  
 PT for treating disorders of central nervous system, pituitary and  
 PT pancreas, and for drug screening  
 XX Diaclosure; Page 134; 206pp; English.

XX This represents the matured murine pituitary-derived ligand polypeptide  
 CC sequence. The polypeptide is a ligand for the G-protein coupled orphan  
 CC receptor designated GPR10 (human) or UHR-1 (rat). Cells transformed with  
 CC a vector containing the ligand polypeptide encoding DNA are used to  
 CC produce a recombinant ligand polypeptide. The ligand polypeptide, and its  
 CC fragments, modulate function of the pituitary, central nervous system,  
 CC pancreas and other tissues and can be used to screen for agents that  
 CC modulate binding of the polypeptide to the receptor; to quantify the  
 CC amount of receptor in a sample and to raise antibodies. They may also be  
 CC used therapeutically, e.g. to treat senile dementia; Alzheimer's,  
 CC Parkinson's or Huntington's diseases; Creutzfeld-Jakob disease; poisoning  
 CC by heavy metals or drugs; diabetes; schizophrenia; disorders of growth  
 CC hormone secretion; cancer; rheumatoid arthritis; epilepsy and many  
 CC others, also to improve post-operative nutritional status and as  
 CC vasopressor. Transgenic animals carrying the ligand polypeptide encoding  
 CC DNA or its mutant are used to study the function of the polypeptide-  
 CC expressing genes, as models of disease, for drug screening and as source  
 CC of cell lines. The ligand polypeptide DNA is used as a source of probes  
 CC and primers; to identify related sequences; in receptor-binding assays;  
 CC for production of Ab and antisera; in drug development; for gene therapy  
 CC and to develop transgenic animals.

XX Sequence 31 AA;

Query Match 100.0%; Score 174; DB 20; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-18;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQSMETRTPDINPAWYTGIRPVGRF 31  
 |||||  
 Db 1 SRAHQSMETRTPDINPAWYTGIRPVGRF 31

## RESULT 5

AAW95174  
 ID AAW95174 standard; Protein; 31 AA.

XX  
 AC AAW95174;

XX 10-MAR-1999 (first entry)

XX Murine pituitary-derived ligand polypeptide antigenic epitope.

XX Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;  
 KW GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas;  
 KW tissue; screen; therapeutic; binding; senile dementia; ligand; murine;  
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;  
 KW Creutzfeld-Jakob disease; poisoning; schizophrenia; growth hormone;  
 KW secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;  
 KW gene therapy; transgenic animal; epitope.

XX Mus sp.

XX WO9849295-A1.

XX 05-NOV-1998.

XX 27-APR-1998; 98WO-JP01923.

XX 28-APR-1997; 97JP-0109974.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Fukusumi S, Hiruma S;

XX

DR WPI; 1999-009423/01.  
 XX New polypeptide ligand for orphan G protein coupled receptors - used  
 PT for treating disorders of central nervous system, pituitary and  
 PT pancreas, and for drug screening  
 XX  
 PS Disclosure; Page 26; 206pp; English.  
 XX  
 CC The invention relates to a murine pituitary-derived ligand polypeptide  
 CC which is a ligand for the G-protein coupled orphan receptor designated  
 CC GPR10 (human) or UHR-1 (rat). Cells transformed with a vector containing  
 CC the ligand polypeptide encoding DNA are used to produce a recombinant  
 CC ligand polypeptide. The ligand polypeptide, and its fragments, modulate  
 CC function of the pituitary, central nervous system, pancreas and other  
 CC tissues and can be used to screen for agents that modulate binding of  
 CC the polypeptide to the receptor; to quantify the amount of receptor in a  
 CC sample and to raise antibodies. They may also be used therapeutically,  
 CC e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's  
 CC diseases; Creutzfeld-Jakob disease; poisoning by heavy metals or drugs;  
 CC diabetes; achizophrenia; disorders of growth hormone secretion; cancer;  
 CC rheumatoid arthritis; epilepsy and many others, also to improve post-  
 CC operative nutritional status and as vasopressor. Transgenic animals  
 CC carrying the ligand polypeptide encoding DNA or its mutin are used to  
 CC study the function of the polypeptide-expressing genes, as models of  
 CC disease, for drug screening and as source of cell lines. The ligand  
 CC polypeptide DNA is used as a source of probes and primers; to identify  
 CC related sequences; in receptor-binding assays; for production of Ab and  
 CC antisera; in drug development; for gene therapy to AAW95178 represent antigenic  
 CC epitopes which can be used for the preparation of anti-ligand polypeptide  
 CC antibody.  
 XX  
 SQ Sequence 31 AA;  
 Query Match 100.0%; Score 174; DB 20; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-18;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SRAHQSMETRTDINPAWYTGIRPVGRF 31  
 DB 1 SRAHQSMETRTDINPAWYTGIRPVGRF 31  
 RESULT 6  
 AAB10355  
 ID AAB10355 standard; peptide; 31 AA.  
 AC AAB10355;  
 XX  
 DT 24-NOV-2000 (first entry)  
 XX  
 DE Rat oxytocin secretion promoting peptide SEQ ID NO: 18.  
 XX  
 KW Rat; oxytocin secretion promoter; G protein-coupled receptor protein;  
 KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
 KW caesarean section; artificial fertilization; galactostasis; goat; pig;  
 KW veterinary medicine; milk production.  
 XX  
 OS Rattus sp.  
 PN WO200038704-A1.  
 PD 06-JUL-2000.  
 XX  
 PF 22-DEC-1999; 99WO-JP07199.  
 XX  
 PR 25-DEC-1998; 98JP-0369585.  
 XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PI Matsumoto H, Kitada C, Hinuma S;  
 XX  
 XX WPI; 2000-452298/39.

XX Physiologically-active polypeptide recognized as ligand by G  
 PT protein-coupled receptor protein, for promoting secretion of oxytocin,  
 PT as drugs for diseases relating to oxytocin secretion and in veterinary  
 PT medicine -  
 XX  
 PS Claim 3; Page 57; 72pp; Japanese.  
 XX  
 CC This invention describes a novel oxytocin secretion-regulating agent  
 CC which contains a ligand peptide or its salt for the G protein-coupled  
 CC receptor protein. It is useful in the form of drugs for ameliorating,  
 CC preventing and treating diseases relating to oxytocin secretion e.g.  
 CC weak pains and atonic bleeding, before and after expulsion of placenta,  
 CC uterine recovery failure, caesarean section, stoppage of artificial  
 CC fertilization or galactostasis and is also applicable in veterinary  
 CC medicine for promoting milk production in cow, goat and pig. This  
 CC sequence represents a rat peptide which acts as an oxytocin secretion  
 CC promoter.  
 XX  
 SQ Sequence 31 AA;  
 Query Match 100.0%; Score 174; DB 21; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-18;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SRAHQSMETRTDINPAWYTGIRPVGRF 31  
 DB 1 SRAHQSMETRTDINPAWYTGIRPVGRF 31  
 RESULT 7  
 AAY87504  
 ID AAY87504 standard; protein; 31 AA.  
 XX  
 AC AAY87504;  
 XX  
 DT 18-JUL-2000 (first entry)  
 XX  
 DE Rat prolactin-releasing peptide, PrRP.  
 XX  
 KW Prolactin-releasing peptide; PrRP; GPR10; G protein-coupled receptor;  
 KW feeding behaviour; food intake; modulation; antagonist; anorectic;  
 KW obesity; agonist; cachexia.  
 XX  
 OS Rattus sp.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 31  
 FT /note= "C-terminal amide"  
 XX  
 PN WO200017641-A1.  
 XX  
 PD 30-MAR-2000.  
 XX  
 PF 22-SEP-1999; 99WO-US21243.  
 XX  
 PR 22-SEP-1998; 98US-0101380.  
 PR 14-OCT-1998; 98US-0172353.  
 XX  
 PA (MILL-) MILLENNIUM PHARM INC.  
 XX  
 PI Stricker-Kongrad A, Gu W;  
 XX  
 XX WPI; 2000-303231/26.  
 DR  
 XX Identifying modulators of body weight by a combination of a cell-free  
 PT or cell-based assay to identify modulators of GPR10, followed by an in  
 PT vivo assay for the compounds effect on e.g. feeding behavior -  
 XX  
 PS Example 2; Page 61; 82pp; English.  
 XX  
 CC The invention relates to a method for identifying compounds useful for  
 CC modulating body weight. The method comprises cell-free and/or cell-based

CC assays that identify compounds which bind to and/or activate or inhibit  
 CC the activity of GPR10, a G protein-coupled receptor. These assays are  
 CC then followed by an in vivo assay of the effect of the compound on  
 CC feeding behaviour, body weight or metabolic rate in a mammal. Prolactin-  
 CC releasing peptide (PrRP; AAY87504) is a ligand of GPR10. Binding of PrRP  
 CC to GPR10 stimulates a signal transduction cascade, which results in an  
 CC increase in food intake. Compounds identified using the method of the  
 CC invention are useful for the modulation of body weight. Antagonists of  
 CC GPR10 can be used to treat obesity, while GPR10 agonists can be used to  
 CC treat cachexia. The present sequence represents rat PrRP.

XX SQ Sequence 31 AA;

Query Match 100.0%; Score 174; DB 21; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-18;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRAHQSMETRTDINPAWYTGIRPVGRF 31  
 |||||  
 Db 1 SRAHQSMETRTDINPAWYTGIRPVGRF 31  
 |||||

## RESULT 8

AAY49292  
 ID AAY49292 standard; peptide; 31 AA.

XX AC AAY49292;

XX DT 22-FEB-2000 (first entry)

XX DE 19P2 ligand peptide fragment.

XX KW Monoclonal antibody; 19P2 ligand; diagnosis; prolactin secretion;

XX KW pituitary; regulatory mechanism; central nervous system; pancreatic.

XX OS Rattus sp.

XX FH Key Location/Qualifiers

XX FT Modified-site 31

XX FT /note= "C-terminal amide"

XX PN WO9960112-A1.

XX PD 25-NOV-1999.

XX PF 20-MAY-1999; 99WO-JP02650.

XX PR 21-MAY-1998; 98JP-0140293.

XX PA (TAKE ) TAKEDA CHEM IND LTD.

XX PI Matsumoto H, Kitada C, Hinuma S;

XX DR WPI; 2000-039381/03.

XX PT New monoclonal antibodies, useful in diagnosis, as drugs and in

XX PT studying diseases related to ligand abnormality -

XX PS Disclosure; Page 26; 73pp; Japanese.

XX CC The invention provides a monoclonal antibody which has a specific

XX CC reaction with the part peptide of the C-terminal of 19P2 ligand or its

XX CC derivative. The antibodies can be used in diagnosis or to treat or

XX CC prevent diseases associated with abnormality in the pituitary function

XX CC regulatory mechanism (e.g. promotion of prolactin secretion), central

XX CC nervous regulatory mechanism, and pancreatic function regulatory

XX CC mechanism. The antibody-based immunoassay can also be applied in

XX CC clarifying the physiological functions of the ligand and its derivative.

XX CC Sequences AAY49290-302 represent peptide fragments of the 19P2 ligand.

XX SQ Sequence' 31 AA;

Query Match 100.0%; Score 174; DB 21; Length 31;

Best Local Similarity 100.0%; Pred. No. 1.8e-18;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRAHQSMETRTDINPAWYTGIRPVGRF 31  
 |||||  
 Db 1 SRAHQSMETRTDINPAWYTGIRPVGRF 31  
 |||||

## RESULT 9

AAG62524  
 ID AAG62524 standard; peptide; 31 AA.

XX AC AAG62524;

XX DT 24-AUG-2001 (first entry)

XX DE Rat CRH releasing protein related peptide SEQ ID NO: 18.

XX KW Rat; corticotrophin releasing hormone; CRH; G protein receptor ligand;

XX KW analgesic; hyperaldosteronism; hypercortisolaemia; hypoadrenocorticism;

XX KW Addison's disease; adrenal gland hyperfunction; obesity.

XX OS Rattus sp.

XX PN WO200135984-A1.

XX PD 25-MAY-2001.

XX PF 17-NOV-2000; 2000WO-JP08119.

XX PR 18-NOV-1999; 99JP-0327900.

XX PR 26-SEP-2000; 2000JP-0297073.

XX PA (TAKE ) TAKEDA CHEM IND LTD.

XX PI Kitada C, Matsumoto H, Hinuma S;

XX DR WPI; 2001-355552/37.

XX FT Use of G protein receptor ligand or peptide for controlling

XX FT corticotropin releasing hormone secretion -

XX PS Claim 3; Page 69; 90pp; Japanese.

XX CC The present sequence describes a method of controlling the secretion of

XX CC corticotrophin releasing hormone (CRH), involving the use of a G protein

XX CC receptor ligand. This can be used to control the secretion of CRH and is

XX CC useful as an analgesic or for treating, preventing or ameliorating

XX CC diseases associated with CRH secretion such as hyperaldosteronism,

XX CC hypercortisolaemia, secondary or chronic hypoadrenocorticism, Addison's

XX CC disease (including boredom, nausea, pigmentation, hypogonadism, hair

XX CC loss, and hypotension), adrenal gland hypofunction and obesity. The

XX CC present sequence is a peptide used in the exemplification of the

XX CC invention.

XX SQ Sequence 31 AA;

Query Match 100.0%; Score 174; DB 22; Length 31;

Best Local Similarity 100.0%; Pred. No. 1.8e-18;

Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRAHQSMETRTDINPAWYTGIRPVGRF 31  
 |||||  
 Db 1 SRAHQSMETRTDINPAWYTGIRPVGRF 31  
 |||||

## RESULT 10

AAB90993  
 ID AAB90993 standard; Peptide; 31 AA.

XX AC AAB90993;

XX DT 22-JUN-2001 (first entry)

XX Prolactin releasing peptide SEQ ID NO:167.  
 DE  
 XX  
 XX Protection; endogenous therapeutic peptide; peptidase; conjugation;  
 KW blood component; modification; succinimidyl; maleimido group; amino;  
 KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.  
 XX  
 XX Homo sapiens.  
 OS Synthetic.  
 OS  
 XX WO200069900-A2.  
 XX  
 XX 23-NOV-2000.  
 XX  
 XX 17-MAY-2000; 2000WO-US13576.  
 XX  
 XX 17-MAY-1999; 99US-0134406.  
 PR 10-SEP-1999; 99US-0153406.  
 PR 15-OCT-1999; 99US-0159783.  
 XX  
 XX (CONJ-) CONJUCHEM INC.  
 PA  
 XX Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;  
 PI WPI; 2001-112059/12.  
 XX  
 XX Modifying and attaching therapeutic peptides to albumin prevents  
 PT peptidase degradation, useful for increasing length of in vivo activity  
 PT  
 XX Disclosure; Page 244; 733pp; English.  
 PS  
 XX The present invention describes a modified therapeutic peptide (I)  
 CC comprising a therapeutically active amino acid region (iii) and a  
 CC reactive group (ii) (e.g. succinimidyl and maleimido groups) attached to  
 CC a less therapeutically active amino acid region (iv), which covalently  
 CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
 CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
 CC (i) are useful for modifying therapeutic peptides e.g. hormones, growth  
 CC factors and neurotransmitters, to protect them from peptidase activity  
 CC in vivo for the treatment of various disorders. Endogenous therapeutic  
 CC peptides are not suitable as drug candidates as they require frequent  
 CC administration due to rapid degradation by peptidases in the body.  
 CC Modifying and attaching therapeutic peptides to albumin prevents or  
 CC reduces the action of peptidases to increase length of activity (half  
 CC life) and specificity as bonding to large molecules decreases  
 CC intracellular uptake and interference with physiological processes.  
 CC AAB90829 to AAB92441 represent peptides which can be used in the  
 CC exemplification of the present invention.  
 XX  
 XX SQ Sequence 31 AA;

Query Match 100.0%; Score 174; DB 22; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-18;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSMETRTDINPAWYTGIRPVGRF 31  
 |||||  
 DB 1 SRAHQHSMETRTDINPAWYTGIRPVGRF 31  
 |||||

RESULT 11  
 AAE26400  
 ID AAE26400 standard; peptide; 31 AA.  
 XX  
 AC AAE26400;  
 XX  
 DT 13-DEC-2002 (first entry)  
 XX  
 XX Rat PrRP-31 peptide.  
 XX  
 XX Rat; wakefulness; sleep disorder; prolactin releasing peptide receptor;  
 KW PrRP; GPR10; therapy; epilepsy; narcolepsy; sleepiness; sleep apnoea;  
 KW

KW insomnia; idiopathic hypersomnia; psychogenic hypersomnia; seizure;  
 KW anticonvulsant.  
 XX  
 OS Rattus sp.  
 XX  
 PN US2002037533-A1.  
 XX  
 XX 28-MAR-2002.  
 XX  
 XX 17-AUG-2001; 2001US-0932161.  
 XX  
 XX 28-APR-2000; 2000US-0560915.  
 XX  
 XX (CIVE/) CIVELLI O.  
 PA (LINS/) LIN S.  
 XX  
 XX Civelli O, Lin S;  
 XX  
 XX WPI; 2002-403931/43.  
 DR  
 XX  
 XX Screening for compounds useful for promoting wakefulness or sleep, and  
 PT for treating sleeping disorders, e.g. insomnia, hypersomnia or sleep  
 PT apnea, comprises administering a prolactin releasing peptide agonist or  
 PT antagonist  
 XX  
 XX Disclosure; Page 24; 35pp; English.  
 PS  
 XX The present invention relates to a method of screening for compounds  
 CC for promoting wakefulness or sleep in a mammal. The method involves  
 CC administering a prolactin releasing peptide (PrRP) receptor (GPR10)  
 CC agonist or antagonist respectively and determining the ability of the  
 CC compound to promote wakefulness or sleep. The compounds identified  
 CC from the method are used in the therapy of epilepsy and other diseases  
 CC associated with absence seizures and in promoting wakefulness and sleep  
 CC in individuals having sleep disorders such as insomnia and narcolepsy.  
 CC PrRP receptor agonists may be used to treat common disorders which lead  
 CC to sleepiness, e.g. sleep apnoea, narcolepsy, idiopathic hypersomnia  
 CC and psychogenic hypersomnia. PrRP receptor antagonists are useful for  
 CC promoting sleep and for treating insomnia such as adjustment sleep  
 CC disorder and psychophysiological insomnia. The present sequence is rat  
 CC PrRP-31 peptide.  
 XX  
 XX SQ Sequence 31 AA;

Query Match 100.0%; Score 174; DB 23; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-18;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSMETRTDINPAWYTGIRPVGRF 31  
 |||||  
 DB 1 SRAHQHSMETRTDINPAWYTGIRPVGRF 31  
 |||||

RESULT 12  
 ABU60826  
 ID ABU60826 standard; Peptide; 31 AA.  
 XX  
 AC ABU60826;  
 XX  
 DT 06-MAY-2003 (first entry)  
 XX  
 XX Peptide production by gene recombination associated peptide #10.  
 DE  
 XX Peptide production; low-molecular peptide; KISS-1; GPR8 ligand;  
 KW gene recombination.  
 KW  
 XX Rattus sp.  
 OS  
 XX WO200292829-A1.  
 PN  
 XX 21-NOV-2002.  
 PD  
 XX 16-MAY-2002; 2002WO-JP04735.  
 PP



XX 17-MAY-2001; 2001JP-0147341.  
 XX (TAKE ) TAKEDA CHEM IND LTD.  
 XX Nishimura O, Suenaga M, Ito T, Kitada C;  
 XX WPI; 2003-129302/12.  
 XX Process for producing peptides e.g. KiSS-1 peptide and GPR8 ligand for  
 PT subsequent applications by gene recombination technique through tandem  
 PT repeats to provide precursor protein with specific cleavage sites -  
 XX Disclosure; Page 58; 87pp; Japanese.  
 XX The invention describes a method of producing a peptide comprising the  
 CC excision of the N and C-terminals of a target peptide with enzymes or  
 CC chemically through the attached cleavage sites repeated by ligation in a  
 CC precursor protein. The method is for producing (low-molecular) peptides  
 CC e.g. KiSS-1 peptide and GPR8 ligand for subsequent applications by the  
 CC gene recombination technique through tandem repeats to provide  
 CC a precursor protein with specific cleavage sites. With this method,  
 CC peptide production can be carried out easily to provide large quantities  
 CC of the required peptides. This is the amino acid sequence of a peptide  
 CC associated with the peptide production method of the invention.  
 XX Sequence 31 AA;  
 SQ Query Match 100.0%; Score 174; DB 24; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-18;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SRAHQSHMETRTPDINPAWYTGIRPVGRF 31  
 DB 1 SRAHQSHMETRTPDINPAWYTGIRPVGRF 31  
 RESULT 13  
 ABU60837  
 ID ABU60837 standard; Peptide; 31 AA.  
 XX AC ABU60837;  
 XX 06-MAY-2003 (first entry)  
 XX Peptide production by gene recombination associated peptide #21.  
 XX Peptide production; low-molecular peptide; KiSS-1; GPR8 ligand;  
 XX gene recombination.  
 XX Rattus sp.  
 XX WO200292829-A1.  
 XX 21-NOV-2002.  
 XX 16-MAY-2002; 2002WO-JP04735.  
 XX 17-MAY-2001; 2001JP-0147341.  
 XX (TAKE ) TAKEDA CHEM IND LTD.  
 XX Nishimura O, Suenaga M, Ito T, Kitada C;  
 XX WPI; 2003-129302/12.  
 XX Process for producing peptides e.g. KiSS-1 peptide and GPR8 ligand for  
 PT subsequent applications by gene recombination technique through tandem  
 PT repeats to provide precursor protein with specific cleavage sites -  
 XX Disclosure; Page 64; 87pp; Japanese.  
 XX The invention describes a method of producing a peptide comprising the

CC excision of the N and C-terminals of a target peptide with enzymes or  
 CC chemically through the attached cleavage sites repeated by ligation in a  
 CC precursor protein. The method is for producing (low-molecular) peptides  
 CC e.g. KiSS-1 peptide and GPR8 ligand for subsequent applications by the  
 CC gene recombination technique through tandem repeats to provide  
 CC a precursor protein with specific cleavage sites. With this method,  
 CC peptide production can be carried out easily to provide large quantities  
 CC of the required peptides. This is the amino acid sequence of a peptide  
 CC associated with the peptide production method of the invention.  
 XX Sequence 31 AA;  
 SQ Query Match 100.0%; Score 174; DB 24; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-18;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SRAHQSHMETRTPDINPAWYTGIRPVGRF 31  
 DB 1 SRAHQSHMETRTPDINPAWYTGIRPVGRF 31  
 RESULT 14  
 AAW31385  
 ID AAW31385 standard; Peptide; 32 AA.  
 XX AC AAW31385;  
 XX 06-APR-1998 (first entry)  
 XX Rat type G protein-coupled receptor ligand fragment 2.  
 XX G protein-coupled receptor; ligand binding; pharmaceutical;  
 XX modulator; pituitary; central nervous system; pancreas; prophylactic;  
 XX therapeutic agent.  
 XX Rat sp.  
 XX WO9724436-A2.  
 XX 10-JUL-1997.  
 XX 26-DEC-1996; 96WO-JP03821.  
 XX 18-SEP-1996; 96JP-0246573.  
 XX 28-DEC-1995; 95JP-0343371.  
 XX 15-MAR-1996; 96JP-0059419.  
 XX 12-AUG-1996; 96JP-0211805.  
 XX (TAKE ) TAKEDA CHEM IND LTD.  
 XX Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;  
 XX Kawamata Y, Kitada C;  
 XX WPI; 1997-363672/33.  
 XX N-PSDB; AAV02422.  
 XX Ligand peptide for G protein-coupled receptor - acts by modulating  
 PT function in the central nervous system, pancreas and pituitary gland  
 XX Claim 2; Page 179; 258pp; English.  
 XX This sequence represents a peptide fragment from a novel rat type  
 CC ligand polypeptide corresponding to amino acid residues 22 to 53 of the  
 CC sequence represented in AAW31383 and is used in an assay to monitor  
 CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
 CC compositions containing this ligand may be used as a pituitary function  
 CC modulator, a central nervous system modulator or a pancreatic function  
 CC modulator. This ligand could have specific applications as a  
 CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
 CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
 CC trauma, growth hormone secretory disease, hyper- and polyphagia,  
 CC hyperlipidaemia, hypercholesterolaemia, hyperglycaemia,  
 CC hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease,

CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
 CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
 CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
 CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
 CC oligosaccharia. Assays can also be developed to screen compounds which are  
 CC capable of altering the binding activity of the ligand affecting  
 CC activation of the G protein-coupled receptor protein.

XX  
 SQ Sequence 32 AA;

Query Match 100.0%; Score 174; DB 18; Length 32;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-18;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQSMETRTDPINPAWYTGIRPVGRF 31  
 |||||  
 DB 1 SRAHQSMETRTDPINPAWYTGIRPVGRF 31

## RESULT 15

AAB10356  
 ID AAB10356 standard; peptide; 32 AA.

XX  
 AC AAB10356;

XX 24-NOV-2000 (first entry)

XX Rat oxytocin secretion promoting peptide SEQ ID NO: 19.

XX Rat; oxytocin secretion promoter; G protein-coupled receptor protein;  
 KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
 KW caesarean section; artificial fertilization; galactostasis; goat; pig;  
 KW veterinary medicine; milk production.

XX Rattus sp.

XX WO200038704-A1.

XX 06-JUL-2000.

XX 22-DEC-1999; 99WO-JP07199.

XX 25-DEC-1998; 98JP-0369585.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Matsumoto H, Kitada C, Hinuma S;

XX WPI; 2000-452298/39.

XX Physiologically-active polypeptide recognized as ligand by G  
 PT protein-coupled receptor protein, for promoting secretion of oxytocin,  
 PT as drugs for diseases relating to oxytocin secretion and in veterinary  
 PT medicine -

XX Disclosure; Page 57; 72pp; Japanese.

XX This invention describes a novel oxytocin secretion-regulating agent  
 CC which contains a ligand peptide or its salt for the G protein-coupled  
 CC receptor protein. It is useful in the form of drugs for ameliorating,  
 CC preventing and treating diseases relating to oxytocin secretion e.g.  
 CC weak pains and atonic bleeding, before and after expulsion of placenta,  
 CC uterine recovery failure, caesarean section, stoppage of artificial  
 CC fertilization or galactostasis and is also applicable in veterinary  
 CC medicine for promoting milk production in cow, goat and pig. This  
 CC sequence represents a rat peptide which acts as an oxytocin secretion  
 CC promoter.

XX Sequence 32 AA;

Query Match 100.0%; Score 174; DB 21; Length 32;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-18;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQSMETRTDPINPAWYTGIRPVGRF 31  
 |||||  
 DB 1 SRAHQSMETRTDPINPAWYTGIRPVGRF 31

Search completed: December 3, 2003, 19:11:24  
 Job time : 33.75 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: December 3, 2003, 19:09:05 ; Search time 11.5 Seconds  
(without alignments)  
114.055 Million cell updates/sec

Title: US-09-868-885B-18

Perfect score: 174

Sequence: 1 SRAHQHSMETRTDINPAWYTGIRPVGRF 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/ias/5A COMB.pcp.\*  
2: /cgn2\_6/ptodata/1/ias/5B COMB.pcp.\*  
3: /cgn2\_6/ptodata/1/ias/6A COMB.pcp.\*  
4: /cgn2\_6/ptodata/1/ias/6B COMB.pcp.\*  
5: /cgn2\_6/ptodata/1/ias/PCTUS COMB.pcp.\*  
6: /cgn2\_6/ptodata/1/ias/backfiles1.pcp.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description       |
|------------|-------|-------------|--------|-------|-------------------|
| 1          | 174   | 100.0       | 31     | 3     | US-09-105-678A-8  |
| 2          | 174   | 100.0       | 31     | 3     | US-09-105-678A-37 |
| 3          | 174   | 100.0       | 31     | 3     | US-09-172-353-4   |
| 4          | 174   | 100.0       | 31     | 3     | US-08-776-971-47  |
| 5          | 174   | 100.0       | 31     | 3     | US-09-421-208-8   |
| 6          | 174   | 100.0       | 31     | 3     | US-09-421-208-37  |
| 7          | 174   | 100.0       | 31     | 4     | US-09-560-915-14  |
| 8          | 174   | 100.0       | 31     | 4     | US-09-799-955-4   |
| 9          | 174   | 100.0       | 32     | 3     | US-09-105-678A-38 |
| 10         | 174   | 100.0       | 32     | 3     | US-08-776-971-48  |
| 11         | 174   | 100.0       | 32     | 3     | US-09-421-208-38  |
| 12         | 174   | 100.0       | 33     | 3     | US-09-105-678A-39 |
| 13         | 174   | 100.0       | 33     | 3     | US-08-776-971-49  |
| 14         | 174   | 100.0       | 33     | 3     | US-09-421-208-39  |
| 15         | 174   | 100.0       | 83     | 3     | US-08-776-971-45  |
| 16         | 174   | 100.0       | 83     | 3     | US-08-776-971-124 |
| 17         | 174   | 100.0       | 83     | 3     | US-08-776-971-137 |
| 18         | 163   | 93.7        | 31     | 3     | US-09-105-678A-7  |
| 19         | 163   | 93.7        | 31     | 3     | US-09-105-678A-31 |
| 20         | 163   | 93.7        | 31     | 3     | US-08-776-971-5   |
| 21         | 163   | 93.7        | 31     | 3     | US-08-776-971-97  |
| 22         | 163   | 93.7        | 31     | 3     | US-09-421-208-7   |
| 23         | 163   | 93.7        | 31     | 3     | US-09-421-208-31  |
| 24         | 163   | 93.7        | 31     | 4     | US-09-560-915-13  |
| 25         | 163   | 93.7        | 32     | 3     | US-09-105-678A-32 |
| 26         | 163   | 93.7        | 32     | 3     | US-08-776-971-6   |
| 27         | 163   | 93.7        | 32     | 3     | US-09-421-208-32  |

Sequence 33, Appli  
Sequence 7, Appli  
Sequence 33, Appli  
Sequence 1, Appli  
Sequence 44, Appli  
Sequence 115, App  
Sequence 117, App  
Sequence 122, App  
Sequence 131, App  
Sequence 136, App  
Sequence 29, Appli  
Sequence 3, Appli  
Sequence 29, Appli  
Sequence 9, Appli  
Sequence 43, Appli  
Sequence 61, Appli  
Sequence 9, Appli  
Sequence 43, Appli

28 163 93.7 33 3 US-09-105-678A-33  
29 163 93.7 33 3 US-08-776-971-7  
30 163 93.7 33 3 US-09-421-208-33  
31 163 93.7 98 3 US-08-776-971-1  
32 163 93.7 98 3 US-08-776-971-44  
33 163 93.7 98 3 US-08-776-971-115  
34 163 93.7 98 3 US-08-776-971-117  
35 163 93.7 98 3 US-08-776-971-122  
36 163 93.7 98 3 US-08-776-971-131  
37 163 93.7 98 3 US-08-776-971-136  
38 152 87.4 29 3 US-09-105-678A-29  
39 152 87.4 29 3 US-08-776-971-3  
40 149 85.6 31 3 US-09-421-208-29  
41 149 85.6 31 3 US-09-105-678A-9  
42 149 85.6 31 3 US-09-105-678A-43  
43 149 85.6 31 3 US-08-776-971-61  
44 149 85.6 31 3 US-09-421-208-9  
45 149 85.6 31 3 US-09-421-208-43

## ALIGNMENTS

## RESULT 1

US-09-105-678A-8  
; Sequence 8, Application US/09105678A  
; Patent No. 6103882  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/105,678A  
; FILING DATE: 26-JUN-1998  
; PRIOR APPLICATION DATA: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 31 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-105-678A-8

Query Match 100.0%; Score 174; DB 3; Length 31;  
Best Local Similarity 100.0%; Pred. No. 1.1e-19;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSMETRTDINPAWYTGIRPVGRF 31

Db 1 SRAHQSMETRTDINPAWYTGIRPVGRF 31  
|||||  
RESULT 2  
US-09-105-678A-37  
; Sequence 37, Application US/09105678A  
; Patent No. 6103882  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/105,678A  
; FILING DATE: 26-JUN-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-8440  
; INFORMATION FOR SEQ ID NO: 37:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 31 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-105-678A-37  
Query Match 100.0%; Score 174; DB 3; Length 31;  
Best Local Similarity 100.0%; Pred. No. 1.1e-19;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SRAHQSMETRTDINPAWYTGIRPVGRF 31  
Db 1 SRAHQSMETRTDINPAWYTGIRPVGRF 31  
|||||  
RESULT 3  
US-09-172-353-4  
; Sequence 4, Application US/09172353  
; Patent No. 6197530  
; GENERAL INFORMATION:  
; APPLICANT: Stricker-Kongra, Alain  
; APPLICANT: Gu, Wei  
; TITLE OF INVENTION: GPRI0 AS A TARGET FOR IDENTIFYING WEIGHT MODULATING COMPOUNDS  
; FILE REFERENCE: 07334/102001  
; CURRENT APPLICATION NUMBER: US/09/172,353  
; CURRENT FILING DATE: 1998-10-14  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 4  
; LENGTH: 31

TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-172-353-4  
Query Match 100.0%; Score 174; DB 3; Length 31;  
Best Local Similarity 100.0%; Pred. No. 1.1e-19;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SRAHQSMETRTDINPAWYTGIRPVGRF 31  
Db 1 SRAHQSMETRTDINPAWYTGIRPVGRF 31  
|||||  
RESULT 4  
US-08-776-971-47  
; Sequence 47, Application US/08776971B  
; Patent No. 6228384  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; ADDRESSEE: Habata, Yugo  
; ADDRESSEE: Kawamata, Yuji  
; ADDRESSEE: Hosoya, Masaki  
; ADDRESSEE: Fujii, Ryo  
; ADDRESSEE: Fukusumi, Shoji  
; ADDRESSEE: Kitada, Chieko  
; TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
; NUMBER OF SEQUENCES: 140  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/776,971B  
; FILING DATE: 06-Feb-1997  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/JP96/03821  
; FILING DATE: 28-DEC-1996  
; APPLICATION NUMBER: JP 7/343371  
; FILING DATE: 28-DEC-1995  
; APPLICATION NUMBER: JP 8/59419  
; FILING DATE: 15-MAR-1996  
; APPLICATION NUMBER: JP 8/211805  
; FILING DATE: 12-AUG-1996  
; APPLICATION NUMBER: JP 8/246573  
; FILING DATE: 18-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 47176  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 47:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 31 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: internal  
; SEQUENCE DESCRIPTION: SEQ ID NO: 47:  
US-08-776-971-47  
Query Match 100.0%; Score 174; DB 3; Length 31;

Best Local Similarity 100.0%; Pred. No. 1.1e-19;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRAHQHSMETRTDINPAWYTGIRPVGRF 31  
Db 1 SRAHQHSMETRTDINPAWYTGIRPVGRF 31

RESULT 5  
US-09-421-208-8  
; Sequence 8, Application US/09421208  
; Patent No. 6258561  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA: US/09/421,208  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/105,678  
; FILING DATE: 26-JUN-1998  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 31 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-421-208-8

Query Match 100.0%; Score 174; DB 3; Length 31;  
Best Local Similarity 100.0%; Pred. No. 1.1e-19;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRAHQHSMETRTDINPAWYTGIRPVGRF 31  
Db 1 SRAHQHSMETRTDINPAWYTGIRPVGRF 31

RESULT 6  
US-09-421-208-37  
; Sequence 37, Application US/09421208  
; Patent No. 6258561  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu

; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA: US/09/421,208  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/105,678  
; FILING DATE: 26-JUN-1998  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 37:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 31 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-421-208-37

Query Match 100.0%; Score 174; DB 3; Length 31;  
Best Local Similarity 100.0%; Pred. No. 1.1e-19;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRAHQHSMETRTDINPAWYTGIRPVGRF 31  
Db 1 SRAHQHSMETRTDINPAWYTGIRPVGRF 31

RESULT 7  
US-09-560-915-14  
; Sequence 14, Application US/09560915  
; Patent No. 6383764  
; GENERAL INFORMATION:  
; APPLICANT: Civeilli, Olivier  
; APPLICANT: Lin, Steven  
; TITLE OF INVENTION: Therapeutic Compositions and Methods  
; TITLE OF INVENTION: Relating To Prolactin Releasing Peptide (PrRP)  
; FILE REFERENCE: P-UC 3534  
; CURRENT APPLICATION NUMBER: US/09/560,915  
; CURRENT FILING DATE: 2000-04-28  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14  
; LENGTH: 31  
; TYPE: PRT  
; ORGANISM: Rattus  
US-09-560-915-14

Query Match 100.0%; Score 174; DB 4; Length 31;  
Best Local Similarity 100.0%; Pred. No. 1.1e-19;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRAHQHSMETRTDINPAWYTGIRPVGRF 31

Db 1 SRAHQHSMETRTDINPANYTGRGIRPVGRF 31

RESULT 8

US-09-799-955-4

Query Match 100.0%; Score 174; DB 4; Length 31;  
Best Local Similarity 100.0%; Pred. No. 1.1e-19;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GENERAL INFORMATION:

APPLICANT: Stricker-Kongra, Alain

APPLICANT: Gu, Wei

TITLE OF INVENTION: GPR10 AS A TARGET FOR IDENTIFYING WEIGHT MODULATING COMPOUNDS

FILE REFERENCE: 07334/102001

CURRENT APPLICATION NUMBER: US/09/799,955

CURRENT FILING DATE: 2001-03-06

PRIOR APPLICATION NUMBER: US/09/172,353

PRIOR FILING DATE: 1998-10-14

NUMBER OF SEQ ID NOS: 7

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 4

LENGTH: 31

TYPE: PRT

ORGANISM: Mus musculus

US-09-799-955-4

Query Match 100.0%; Score 174; DB 4; Length 31;  
Best Local Similarity 100.0%; Pred. No. 1.1e-19;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GENERAL INFORMATION:

APPLICANT: Suenaga, Masato

APPLICANT: Moriya, Takeo

APPLICANT: Tanaka, Yoko

APPLICANT: Nishimura, Osamu

TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND

NUMBER OF SEQUENCES: 52

CORRESPONDENCE ADDRESS:

ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

STREET: 130 Water Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/105,678A

FILING DATE: 26-JUN-1998

PRIOR APPLICATION NUMBER: JP 172118/1997

FILING DATE: 27-JUN-1997

ATTORNEY/AGENT INFORMATION:

NAME: Conlin, David G.

REGISTRATION NUMBER: 27,026

REFERENCE/DOCKET NUMBER: 48466-342

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 38:

SEQUENCE CHARACTERISTICS:

LENGTH: 32 amino acids

TYPE: amino acid

Db 1 SRAHQHSMETRTDINPANYTGRGIRPVGRF 31

RESULT 9

US-09-105-678A-38

Query Match 100.0%; Score 174; DB 4; Length 31;  
Best Local Similarity 100.0%; Pred. No. 1.1e-19;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GENERAL INFORMATION:

APPLICANT: Suenaga, Masato

APPLICANT: Moriya, Takeo

APPLICANT: Tanaka, Yoko

APPLICANT: Nishimura, Osamu

TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND

NUMBER OF SEQUENCES: 52

CORRESPONDENCE ADDRESS:

ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

STREET: 130 Water Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/105,678A

FILING DATE: 26-JUN-1998

PRIOR APPLICATION NUMBER: JP 172118/1997

FILING DATE: 27-JUN-1997

ATTORNEY/AGENT INFORMATION:

NAME: Conlin, David G.

REGISTRATION NUMBER: 27,026

REFERENCE/DOCKET NUMBER: 48466-342

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 38:

SEQUENCE CHARACTERISTICS:

LENGTH: 32 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-09-105-678A-38

Query Match 100.0%; Score 174; DB 3; Length 32;  
Best Local Similarity 100.0%; Pred. No. 1.1e-19;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSMETRTDINPANYTGRGIRPVGRF 31

Db 1 SRAHQHSMETRTDINPANYTGRGIRPVGRF 31

RESULT 10

US-08-776-971-48

Query Match 100.0%; Score 174; DB 3; Length 32;  
Best Local Similarity 100.0%; Pred. No. 1.1e-19;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GENERAL INFORMATION:

APPLICANT: Hinuma, Shuji

APPLICANT: Habata, Yugo

APPLICANT: Kawamata, Yuji

APPLICANT: Hosoya, Masaki

APPLICANT: Fujii, Ryo

APPLICANT: Fukusumi, Shoji

APPLICANT: Kitada, Chieko

TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE

NUMBER OF SEQUENCES: 140

CORRESPONDENCE ADDRESS:

ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

STREET: 130 Water Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/776,971B

FILING DATE: 06-Feb-1997

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/JP96/03821

FILING DATE: 28-DEC-1996

APPLICATION NUMBER: JP 7/343371

FILING DATE: 28-DEC-1995

APPLICATION NUMBER: JP 8/59419

FILING DATE: 15-MAR-1996

APPLICATION NUMBER: JP 8/211805

APPLICATION NUMBER: JP 8/246573

FILING DATE: 18-SEP-1996

ATTORNEY/AGENT INFORMATION:

NAME: Conlin, David G.

REGISTRATION NUMBER: 27,026

REFERENCE/DOCKET NUMBER: 47176

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 48:

SEQUENCE CHARACTERISTICS:

LENGTH: 32 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

FRAGMENT TYPE: internal

SEQUENCE DESCRIPTION: SEQ ID NO: 48:

US-08-776-971-48

```

Query Match      100.0%; Score 174; DB 3; Length 32;
Best Local Similarity 100.0%; Pred. No. 1.1e-19;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRAHQSHMETRTPDINPAWYTGIRPVGRF 31
Db 1 SRAHQSHMETRTPDINPAWYTGIRPVGRF 31

RESULT 11
US-09-421-208-38
; Sequence 38, Application US/09421208
; Patent No. 6258561
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/421.208
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/105.678
; FILING DATE: 26-JUN-1998
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 38:
; LENGTH: 32 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-421-208-38

Query Match      100.0%; Score 174; DB 3; Length 32;
Best Local Similarity 100.0%; Pred. No. 1.1e-19;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRAHQSHMETRTPDINPAWYTGIRPVGRF 31
Db 1 SRAHQSHMETRTPDINPAWYTGIRPVGRF 31

RESULT 12
US-09-105-678A-39
; Sequence 39, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko

```

```

; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105.678A
; FILING DATE: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 39:
; LENGTH: 33 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-105-678A-39

Query Match      100.0%; Score 174; DB 3; Length 33;
Best Local Similarity 100.0%; Pred. No. 1.1e-19;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRAHQSHMETRTPDINPAWYTGIRPVGRF 31
Db 1 SRAHQSHMETRTPDINPAWYTGIRPVGRF 31

RESULT 13
US-08-776-971-49
; Sequence 49, Application US/08776971B
; Patent No. 6228984
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Habata, Yugo
; APPLICANT: Kawamata, Yuji
; APPLICANT: Hosoya, Masaki
; APPLICANT: Fujii, Ryo
; APPLICANT: Fukusumi, Shoji
; APPLICANT: Kitada, Chieko
; TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
; NUMBER OF SEQUENCES: 140
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:

```

```
/ APPLICATION NUMBER: US/08/776,971B
/ FILING DATE: 06-Feb-1997
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: PCT/JP96/03821
/ FILING DATE: 28-DEC-1996
/ APPLICATION NUMBER: JP 7/343371
/ FILING DATE: 28-DEC-1995
/ APPLICATION NUMBER: JP 8/59419
/ FILING DATE: 15-MAR-1996
/ APPLICATION NUMBER: JP 8/211805
/ FILING DATE: 12-AUG-1996
/ APPLICATION NUMBER: JP 8/246573
/ FILING DATE: 18-SEP-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Conlin, David G.
/ REGISTRATION NUMBER: 27,026
/ REFERENCE/DOCKET NUMBER: 47176
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617-523-3400
/ TELEFAX: 617-523-6440
/ INFORMATION FOR SEQ ID NO: 49:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 33 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ FRAGMENT TYPE: internal
/ SEQUENCE DESCRIPTION: SEQ ID NO: 49:
US-08-776-971-49

Query Match 100.0%; Score 174; DB 3; Length 33;
Best Local Similarity 100.0%; Pred. No. 1.1e-19;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQSMETRTDINPAWYTGIRPVGRF 31
Db 1 SRAHQSMETRTDINPAWYTGIRPVGRF 31

RESULT 14
US-09-421-208-39
/ Sequence 39, Application US/09421208
/ Patent No. 6258561
/ GENERAL INFORMATION:
/ APPLICANT: Suenaga, Masato
/ APPLICANT: Moriya, Takeo
/ APPLICANT: Tanaka, Yoko
/ APPLICANT: Nishimura, Osamu
/ TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
/ NUMBER OF SEQUENCES: 52
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
/ STREET: 130 Water Street
/ CITY: Boston
/ STATE: MA
/ COUNTRY: USA
/ ZIP: 02109
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent in Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/421,208
/ FILING DATE:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 09/105,678
/ FILING DATE: 26-JUN-1998
/ APPLICATION NUMBER: JP 172118/1997
/ FILING DATE: 27-JUN-1997
/ ATTORNEY/AGENT INFORMATION:

/ APPLICATION NUMBER: US/08/776,971B
/ FILING DATE: 06-Feb-1997
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: PCT/JP96/03821
/ FILING DATE: 28-DEC-1996
/ APPLICATION NUMBER: JP 7/343371
/ FILING DATE: 28-DEC-1995
/ APPLICATION NUMBER: JP 8/59419
/ FILING DATE: 15-MAR-1996
/ APPLICATION NUMBER: JP 8/211805
/ FILING DATE: 12-AUG-1996
/ APPLICATION NUMBER: JP 8/246573
/ FILING DATE: 18-SEP-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Conlin, David G.
/ REGISTRATION NUMBER: 27,026
/ REFERENCE/DOCKET NUMBER: 47176
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617-523-3400
/ TELEFAX: 617-523-6440
/ INFORMATION FOR SEQ ID NO: 45:

/ NAME: Conlin, David G.
/ REGISTRATION NUMBER: 27,026
/ REFERENCE/DOCKET NUMBER: 48466-342
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617-523-3400
/ TELEFAX: 617-523-6440
/ INFORMATION FOR SEQ ID NO: 39:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 33 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
US-09-421-208-39

Query Match 100.0%; Score 174; DB 3; Length 33;
Best Local Similarity 100.0%; Pred. No. 1.1e-19;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQSMETRTDINPAWYTGIRPVGRF 31
Db 1 SRAHQSMETRTDINPAWYTGIRPVGRF 31

RESULT 15
US-08-776-971-45
/ Sequence 45, Application US/08776971B
/ Patent No. 6228984
/ GENERAL INFORMATION:
/ APPLICANT: Hinuma, Shuji
/ APPLICANT: Habata, Yugo
/ APPLICANT: Kawamata, Yuji
/ APPLICANT: Hosoya, Masaki
/ APPLICANT: Fujii, Ryo
/ APPLICANT: Fukusumi, Shoji
/ APPLICANT: Kitada, Chieko
/ TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
/ NUMBER OF SEQUENCES: 140
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
/ STREET: 130 Water Street
/ CITY: Boston
/ STATE: MA
/ COUNTRY: USA
/ ZIP: 02109
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/776,971B
/ FILING DATE: 06-Feb-1997
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: PCT/JP96/03821
/ FILING DATE: 28-DEC-1996
/ APPLICATION NUMBER: JP 7/343371
/ FILING DATE: 28-DEC-1995
/ APPLICATION NUMBER: JP 8/59419
/ FILING DATE: 15-MAR-1996
/ APPLICATION NUMBER: JP 8/211805
/ FILING DATE: 12-AUG-1996
/ APPLICATION NUMBER: JP 8/246573
/ FILING DATE: 18-SEP-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Conlin, David G.
/ REGISTRATION NUMBER: 27,026
/ REFERENCE/DOCKET NUMBER: 47176
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617-523-3400
/ TELEFAX: 617-523-6440
/ INFORMATION FOR SEQ ID NO: 45:
```



```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 83 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 45:
US-08-776-971-45

```

```

Query Match      100.0%; Score 174; DB 3; Length 83;
Best Local Similarity 100.0%; Pred. No. 3.3e-19;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 SRAHSHMETPTDINPAWYTGIRPVGRF 31
Db 22 SRAHSHMETPTDINPAWYTGIRPVGRF 52

```

```

Search completed: December 3, 2003, 19:15:55
Job time : 11.5 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 19:11:30 ; Search time 22.5 Seconds  
(without alignments)  
256.244 Million cell updates/sec

Title: US-09-868-885B-18

Perfect score: 174

Sequence: 1 SRAHQSMETRTDINPAWYTGIRPVGRF 31

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_5/ptodata/1/pubaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description      |
|------------|-------|-------------|--------|-------|------------------|
| 1          | 174   | 100.0       | 31     | 9     | US-09-932-161-14 |
| 2          | 174   | 100.0       | 31     | 12    | US-10-096-777-14 |
| 3          | 174   | 100.0       | 31     | 14    | US-10-044-592-4  |
| 4          | 174   | 100.0       | 31     | 14    | US-10-044-592-5  |
| 5          | 174   | 100.0       | 70     | 14    | US-10-044-592-90 |
| 6          | 174   | 100.0       | 82     | 14    | US-10-044-592-1  |
| 7          | 174   | 100.0       | 86     | 14    | US-10-044-592-96 |
| 8          | 174   | 100.0       | 91     | 14    | US-10-044-592-94 |
| 9          | 163   | 93.7        | 31     | 9     | US-09-932-161-13 |
| 10         | 163   | 93.7        | 31     | 12    | US-10-096-777-13 |
| 11         | 163   | 93.7        | 31     | 14    | US-10-044-592-39 |
| 12         | 163   | 93.7        | 32     | 14    | US-10-044-592-40 |
| 13         | 163   | 93.7        | 33     | 14    | US-10-044-592-41 |
| 14         | 163   | 93.7        | 98     | 14    | US-10-044-592-28 |
| 15         | 163   | 93.7        | 98     | 14    | US-10-044-592-38 |

|    |     |      |     |    |                     |                   |
|----|-----|------|-----|----|---------------------|-------------------|
| 16 | 163 | 93.7 | 98  | 14 | US-10-044-592-82    | Sequence 82, Appl |
| 17 | 163 | 93.7 | 98  | 14 | US-10-044-592-84    | Sequence 84, Appl |
| 18 | 163 | 93.7 | 98  | 14 | US-10-044-592-86    | Sequence 86, Appl |
| 19 | 163 | 93.7 | 98  | 14 | US-10-044-592-88    | Sequence 88, Appl |
| 20 | 152 | 87.4 | 29  | 14 | US-10-044-592-26    | Sequence 26, Appl |
| 21 | 149 | 85.6 | 31  | 9  | US-09-932-161-15    | Sequence 15, Appl |
| 22 | 149 | 85.6 | 31  | 12 | US-10-096-777-15    | Sequence 92, Appl |
| 23 | 149 | 85.6 | 87  | 14 | US-10-044-592-92    | Sequence 15, Appl |
| 24 | 133 | 76.4 | 25  | 14 | US-10-044-592-78    | Sequence 78, Appl |
| 25 | 116 | 66.7 | 20  | 9  | US-09-932-161-17    | Sequence 17, Appl |
| 26 | 116 | 66.7 | 20  | 12 | US-10-096-777-17    | Sequence 17, Appl |
| 27 | 116 | 66.7 | 20  | 14 | US-10-044-592-6     | Sequence 6, Appl  |
| 28 | 111 | 63.8 | 20  | 9  | US-09-932-161-16    | Sequence 16, Appl |
| 29 | 111 | 63.8 | 20  | 12 | US-10-096-777-16    | Sequence 16, Appl |
| 30 | 111 | 63.8 | 20  | 14 | US-10-044-592-42    | Sequence 42, Appl |
| 31 | 111 | 63.8 | 21  | 14 | US-10-044-592-43    | Sequence 43, Appl |
| 32 | 111 | 63.8 | 22  | 14 | US-10-044-592-44    | Sequence 44, Appl |
| 33 | 105 | 60.3 | 19  | 14 | US-10-044-592-27    | Sequence 27, Appl |
| 34 | 105 | 60.3 | 20  | 9  | US-09-932-161-18    | Sequence 18, Appl |
| 35 | 105 | 60.3 | 20  | 12 | US-10-096-777-18    | Sequence 18, Appl |
| 36 | 91  | 52.3 | 40  | 14 | US-10-044-592-80    | Sequence 80, Appl |
| 37 | 62  | 35.6 | 209 | 14 | US-10-108-915-30    | Sequence 30, Appl |
| 38 | 57  | 32.8 | 9   | 14 | US-10-044-592-8     | Sequence 8, Appl  |
| 39 | 56  | 32.2 | 428 | 9  | US-09-820-155-2     | Sequence 2, Appl  |
| 40 | 54  | 31.0 | 10  | 14 | US-10-044-592-9     | Sequence 9, Appl  |
| 41 | 53  | 30.5 | 465 | 12 | US-10-301-822-197   | Sequence 197, App |
| 42 | 53  | 30.5 | 949 | 10 | US-09-738-626-5020  | Sequence 5020, Ap |
| 43 | 51  | 29.3 | 239 | 9  | US-09-810-808-7     | Sequence 7, Appl  |
| 44 | 51  | 29.3 | 428 | 9  | US-09-820-155-4     | Sequence 4, Appl  |
| 45 | 51  | 29.3 | 943 | 9  | US-09-815-242-10994 | Sequence 10994, A |

## ALIGNMENTS

### RESULT 1

US-09-932-161-14  
; Sequence 14, Application US/09932161  
; Patent No. US20020037533A1  
; GENERAL INFORMATION:  
; APPLICANT: Civealli, Olivier  
; APPLICANT: Lin, Steven  
; TITLE OF INVENTION: Screening and Therapeutic Methods For  
; TITLE OF INVENTION: Promoting Wakefulness and Sleep  
; FILE REFERENCE: P-UC 4679  
; CURRENT APPLICATION NUMBER: US/09/932,161  
; CURRENT FILING DATE: 2001-08-17  
; PRIOR APPLICATION NUMBER: US 09/560,915  
; PRIOR FILING DATE: 2000-04-28  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14  
; LENGTH: 31  
; TYPE: PRT  
; ORGANISM: Rattus  
US-09-932-161-14

Query Match 100.0%; Score 174; DB 9; Length 31;  
Best Local Similarity 100.0%; Pred. No. 9.9e-18;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRAHQSMETRTDINPAWYTGIRPVGRF 31

Db 1 SRAHQSMETRTDINPAWYTGIRPVGRF 31

### RESULT 2

US-10-096-777-14  
; Sequence 14, Application US/10096777  
; Publication No. US20030171270A1  
; GENERAL INFORMATION:  
; APPLICANT: Civealli, Olivier  
; APPLICANT: Lin, Steven

; TITLE OF INVENTION: Therapeutic Compositions and Methods  
; FILE REFERENCE: P-UC 3534  
; CURRENT APPLICATION NUMBER: US/10/096,777  
; PRIOR FILING DATE: 2002-03-12  
; PRIOR APPLICATION NUMBER: US/09/560,915  
; PRIOR FILING DATE: 2000-04-28  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14  
; LENGTH: 31  
; TYPE: PRT  
; ORGANISM: Rattus  
US-10-096-777-14

Query Match 100.0%; Score 174; DB 12; Length 31;  
Best Local Similarity 100.0%; Pred. No. 9.9e-18;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSMETRTDINPAWYTGIRPVGRF 31  
DB 1 SRAHQHSMETRTDINPAWYTGIRPVGRF 31

## RESULT 3

US-10-044-592-4  
; Sequence 4, Application US/10044592  
; Publication No. US20020143152A1  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; TITLE OF INVENTION: Polypeptides, their Production and Use  
; FILE REFERENCE: 2463US2P  
; CURRENT APPLICATION NUMBER: US/10/044,592  
; PRIOR FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: US 09/403639  
; PRIOR FILING DATE: 1999-25-10  
; PRIOR APPLICATION NUMBER: PCT/JP98/01923  
; PRIOR FILING DATE: 1998-04-27  
; PRIOR APPLICATION NUMBER: JP 9-109974  
; PRIOR FILING DATE: 1997-04-28  
; NUMBER OF SEQ ID NOS: 96  
; SOFTWARE:  
; SEQ ID NO 4  
; LENGTH: 31  
; TYPE: PRT  
; ORGANISM: Murine  
US-10-044-592-4

Query Match 100.0%; Score 174; DB 14; Length 31;  
Best Local Similarity 100.0%; Pred. No. 9.9e-18;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSMETRTDINPAWYTGIRPVGRF 31  
DB 1 SRAHQHSMETRTDINPAWYTGIRPVGRF 31

## RESULT 4

US-10-044-592-5  
; Sequence 5, Application US/10044592  
; Publication No. US20020143152A1  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; TITLE OF INVENTION: Polypeptides, their Production and Use  
; FILE REFERENCE: 2463US2P  
; CURRENT APPLICATION NUMBER: US/10/044,592  
; PRIOR FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: US 09/403639  
; PRIOR FILING DATE: 1999-25-10  
; PRIOR APPLICATION NUMBER: PCT/JP98/01923  
; PRIOR FILING DATE: 1998-04-27

; PRIOR APPLICATION NUMBER: JP 9-109974  
; PRIOR FILING DATE: 1997-04-28  
; NUMBER OF SEQ ID NOS: 96  
; SOFTWARE:  
; SEQ ID NO 5  
; LENGTH: 31  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(31)  
; OTHER INFORMATION: antigen  
US-10-044-592-5

Query Match 100.0%; Score 174; DB 14; Length 31;  
Best Local Similarity 100.0%; Pred. No. 9.9e-18;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSMETRTDINPAWYTGIRPVGRF 31  
DB 1 SRAHQHSMETRTDINPAWYTGIRPVGRF 31

## RESULT 5

US-10-044-592-90  
; Sequence 90, Application US/10044592  
; Publication No. US20020143152A1  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; TITLE OF INVENTION: Polypeptides, their Production and Use  
; FILE REFERENCE: 2463US2P  
; CURRENT APPLICATION NUMBER: US/10/044,592  
; CURRENT FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: US 09/403639  
; PRIOR FILING DATE: 1999-25-10  
; PRIOR APPLICATION NUMBER: PCT/JP98/01923  
; PRIOR FILING DATE: 1998-04-27  
; PRIOR APPLICATION NUMBER: JP 9-109974  
; PRIOR FILING DATE: 1997-04-28  
; NUMBER OF SEQ ID NOS: 96  
; SOFTWARE:  
; SEQ ID NO 90  
; LENGTH: 70  
; TYPE: PRT  
; ORGANISM: Rat  
US-10-044-592-90

Query Match 100.0%; Score 174; DB 14; Length 70;  
Best Local Similarity 100.0%; Pred. No. 2.3e-17;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSMETRTDINPAWYTGIRPVGRF 31  
DB 22 SRAHQHSMETRTDINPAWYTGIRPVGRF 52

## RESULT 6

US-10-044-592-1  
; Sequence 1, Application US/10044592  
; Publication No. US20020143152A1  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; TITLE OF INVENTION: Polypeptides, their Production and Use  
; FILE REFERENCE: 2463US2P  
; CURRENT APPLICATION NUMBER: US/10/044,592  
; CURRENT FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: US 09/403639  
; PRIOR FILING DATE: 1999-25-10  
; PRIOR APPLICATION NUMBER: PCT/JP98/01923  
; PRIOR FILING DATE: 1998-04-27  
; PRIOR APPLICATION NUMBER: JP 9-109974

; PRIOR FILING DATE: 1997-04-28  
; NUMBER OF SEQ ID NOS: 96  
; SOFTWARE:  
; SEQ ID NO 1  
; LENGTH: 82  
; TYPE: PRT  
; ORGANISM: Murine  
US-10-044-592-1

Query Match 100.0%; Score 174; DB 14; Length 82;  
Best Local Similarity 100.0%; Pred. No. 2.8e-17;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRAHQHSMETRTDINPAWYTGIRPVGRF 31  
Db 21 SRAHQHSMETRTDINPAWYTGIRPVGRF 51

RESULT 7  
US-10-044-592-96  
; Sequence 96, Application US/10044592  
; Publication No. US20020143152A1  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; TITLE OF INVENTION: Polypeptides, their Production and Use  
; FILE REFERENCE: 2463US2P  
; CURRENT APPLICATION NUMBER: US/10/044,592  
; CURRENT FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: US 09/403639  
; PRIOR FILING DATE: 1999-25-10  
; PRIOR APPLICATION NUMBER: PCT/JP98/01923  
; PRIOR FILING DATE: 1998-04-27  
; PRIOR APPLICATION NUMBER: JP 9-109974  
; PRIOR FILING DATE: 1997-04-28  
; NUMBER OF SEQ ID NOS: 96  
; SOFTWARE:  
; SEQ ID NO 96  
; LENGTH: 86  
; TYPE: PRT  
; ORGANISM: mammalian  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1224)..(1243)  
; OTHER INFORMATION: Bracket region depicted in FIG 39.  
US-10-044-592-96

Query Match 100.0%; Score 174; DB 14; Length 86;  
Best Local Similarity 100.0%; Pred. No. 2.9e-17;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRAHQHSMETRTDINPAWYTGIRPVGRF 31  
Db 22 SRAHQHSMETRTDINPAWYTGIRPVGRF 52

RESULT 8  
US-10-044-592-94  
; Sequence 94, Application US/10044592  
; Publication No. US20020143152A1  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; TITLE OF INVENTION: Polypeptides, their Production and Use  
; FILE REFERENCE: 2463US2P  
; CURRENT APPLICATION NUMBER: US/10/044,592  
; CURRENT FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: US 09/403639  
; PRIOR FILING DATE: 1999-25-10  
; PRIOR APPLICATION NUMBER: PCT/JP98/01923  
; PRIOR FILING DATE: 1998-04-27  
; PRIOR APPLICATION NUMBER: JP 9-109974  
; PRIOR FILING DATE: 1997-04-28

; NUMBER OF SEQ ID NOS: 96  
; SOFTWARE:  
; SEQ ID NO 94  
; LENGTH: 91  
; TYPE: PRT  
; ORGANISM: Mammalian  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)..(31)  
; OTHER INFORMATION: primer  
; NAME/KEY: misc feature  
; LOCATION: (925)..(955)  
; OTHER INFORMATION: primer  
; NAME/KEY: misc feature  
; LOCATION: (1)..(955)  
; OTHER INFORMATION: Insert fragment of pmGB3 as depicted in FIG 36 and 37.  
US-10-044-592-94

Query Match 100.0%; Score 174; DB 14; Length 91;  
Best Local Similarity 100.0%; Pred. No. 3.1e-17;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRAHQHSMETRTDINPAWYTGIRPVGRF 31  
Db 22 SRAHQHSMETRTDINPAWYTGIRPVGRF 52

RESULT 9  
US-09-932-161-13  
; Sequence 13, Application US/09932161  
; Patent No. US20020037533A1  
; GENERAL INFORMATION:  
; APPLICANT: Civelli, Olivier  
; APPLICANT: Lin, Steven  
; TITLE OF INVENTION: Screening and Therapeutic Methods For  
; TITLE OF INVENTION: Promoting Wakefulness and Sleep  
; FILE REFERENCE: P-UC 4679  
; CURRENT APPLICATION NUMBER: US/09/932,161  
; CURRENT FILING DATE: 2001-08-17  
; PRIOR APPLICATION NUMBER: US 09/560,915  
; PRIOR FILING DATE: 2000-04-28  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13  
; LENGTH: 31  
; TYPE: PRT  
; ORGANISM: Bos taurus  
US-09-932-161-13

Query Match 93.7%; Score 163; DB 9; Length 31;  
Best Local Similarity 93.5%; Pred. No. 3.6e-16;  
Matches 29; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SRAHQHSMETRTDINPAWYTGIRPVGRF 31  
Db 1 SRAHQHSMETRTDINPAWYTGIRPVGRF 31

RESULT 10  
US-10-096-777-13  
; Sequence 13, Application US/10096777  
; Publication No. US20030171270A1  
; GENERAL INFORMATION:  
; APPLICANT: Civelli, Olivier  
; APPLICANT: Lin, Steven  
; TITLE OF INVENTION: Therapeutic Compositions and Methods  
; TITLE OF INVENTION: Relating To Prolactin Releasing Peptide (PrRP)  
; FILE REFERENCE: P-UC 3534  
; CURRENT APPLICATION NUMBER: US/10/096,777  
; CURRENT FILING DATE: 2002-03-12  
; PRIOR APPLICATION NUMBER: US/09/560,915  
; PRIOR FILING DATE: 2000-04-28  
; NUMBER OF SEQ ID NOS: 24

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Bos taurus
US-10-096-777-13

Query Match
Best Local Similarity 93.7%; Score 163; DB 12; Length 31;
Matches 29; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SRAHQHSMETRTDINPAWYTGIRPVGRF 31
Db 1 SRAHQHSMETRTDINPAWYAGRGIRPVGRF 31

RESULT 11
US-10-044-592-39
; Sequence 39, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Fukusumi, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 39
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-39

Query Match
Best Local Similarity 93.7%; Score 163; DB 14; Length 31;
Matches 29; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SRAHQHSMETRTDINPAWYTGIRPVGRF 31
Db 1 SRAHQHSMETRTDINPAWYAGRGIRPVGRF 31

RESULT 12
US-10-044-592-40
; Sequence 40, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Fukusumi, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 40
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-40

Query Match
Best Local Similarity 93.7%; Score 163; DB 13; Length 32;
Matches 29; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SRAHQHSMETRTDINPAWYTGIRPVGRF 31
Db 1 SRAHQHSMETRTDINPAWYAGRGIRPVGRF 31

RESULT 13
US-10-044-592-41
; Sequence 41, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Fukusumi, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 41
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-41

Query Match
Best Local Similarity 93.7%; Score 163; DB 14; Length 33;
Matches 29; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SRAHQHSMETRTDINPAWYTGIRPVGRF 31
Db 1 SRAHQHSMETRTDINPAWYAGRGIRPVGRF 31

RESULT 14
US-10-044-592-28
; Sequence 28, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Fukusumi, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 28
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Murine
US-10-044-592-28

Query Match
Best Local Similarity 93.7%; Score 163; DB 14; Length 98;
Matches 29; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

Qy 1 SRAHQHSMETPTDINPAWYTGIRPVGRF 31  
Db 23 SRAHQHSMETPTDINPAWYTGIRPVGRF 53

## RESULT 15

US-10-044-592-38  
; Sequence 38, Application US/10044592  
; Publication No. US20020143152A1  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; APPLICANT: Fukusumi, Shoji  
; TITLE OF INVENTION: Polypeptides, their Production and Use  
; FILE REFERENCE: 2463US2P  
; CURRENT APPLICATION NUMBER: US/10/044.592  
; CURRENT FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: US 09/403639  
; PRIOR FILING DATE: 1999-25-10  
; PRIOR APPLICATION NUMBER: PCT/JP98/01923  
; PRIOR FILING DATE: 1998-04-27  
; PRIOR APPLICATION NUMBER: JP 9-109974  
; PRIOR FILING DATE: 1997-04-28  
; NUMBER OF SEQ ID NOS: 96  
; SOFTWARE:  
; SEQ ID NO 38  
; LENGTH: 98  
; TYPE: PRT  
; ORGANISM: Bovine  
US-10-044-592-38

Query Match 93.7%; Score 163; DB 14; Length 98;  
Best Local Similarity 93.5%; Pred. No. 1.2e-15;  
Matches 29; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SRAHQHSMETPTDINPAWYTGIRPVGRF 31  
Db 23 SRAHQHSMETPTDINPAWYTGIRPVGRF 53

Search completed: December 3, 2003, 19:17:37  
Job time : 22.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 19:05:30 ; Search time 10.5 Seconds  
(without alignments)  
283.927 Million cell updates/sec

Title: US-09-868-885B-18

Perfect score: 174

Sequence: 1 SRAHQSMETRTDINPAWYTGIRPVGRF 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query % | Length | ID       | Description        |
|------------|-------|---------|--------|----------|--------------------|
| 1          | 174   | 100.0   | 83     | 2 JC7607 | prolactin-releasin |
| 2          | 61    | 35.1    | 1236   | 2 T50904 | Mg protoporphyrin  |
| 3          | 58    | 33.3    | 1292   | 2 T31462 | probable magnesium |
| 4          | 56    | 32.2    | 962    | 2 H69157 | excinuclease ABC c |
| 5          | 53    | 30.5    | 798    | 2 S11210 | probable unr prote |
| 6          | 53    | 30.5    | 1415   | 2 C83070 | conserved hypothet |
| 7          | 52.5  | 30.2    | 303    | 2 AH2016 | hypothetical prote |
| 8          | 52    | 29.9    | 940    | 2 A82329 | excinuclease ABC c |
| 9          | 52    | 29.9    | 969    | 2 B87083 | excinuclease ABC a |
| 10         | 52    | 29.9    | 972    | 2 A70619 | excinuclease ABC c |
| 11         | 52    | 29.9    | 974    | 2 AH3361 | excinuclease ABC c |
| 12         | 51.5  | 29.6    | 503    | 2 A82193 | Sun/nucleolar prot |
| 13         | 51    | 29.3    | 294    | 2 T21075 | hypothetical prote |
| 14         | 51    | 29.3    | 943    | 2 D64057 | excinuclease ABC c |
| 15         | 50.5  | 29.0    | 176    | 2 S67150 | hypothetical prote |
| 16         | 50.5  | 29.0    | 548    | 2 T47548 | hypothetical prote |
| 17         | 50    | 28.7    | 264    | 2 C84971 | hypothetical prote |
| 18         | 50    | 28.7    | 482    | 1 S40887 | RVS167 protein - Y |
| 19         | 50    | 28.7    | 940    | 1 BVECUA | excinuclease ABC c |
| 20         | 50    | 28.7    | 940    | 2 H91258 | excision nuclease  |
| 21         | 50    | 28.7    | 940    | 2 D86099 | excision nuclease  |
| 22         | 50    | 28.7    | 941    | 2 AE1017 | excision nuclease  |
| 23         | 50    | 28.7    | 947    | 2 AF0040 | excinuclease ABC c |
| 24         | 50    | 28.7    | 953    | 2 D71645 | excinuclease ABC c |
| 25         | 50    | 28.7    | 955    | 2 F37861 | excinuclease ABC a |
| 26         | 50    | 28.7    | 965    | 2 C82560 | excinuclease ABC c |
| 27         | 50    | 28.7    | 973    | 2 AH2762 | ABC excinuclease a |
| 28         | 50    | 28.7    | 982    | 2 G37543 | excinuclease ABC c |
| 29         | 49.5  | 28.4    | 375    | 2 F91173 | probable transport |

30 49.5 28.4 375 2 F86019 probable transport  
31 49.5 28.4 375 2 S47704 hypothetical 41.1K  
32 49 28.2 128 2 S76955 hypothetical prote  
33 49 28.2 220 2 C83292 probable glutathio  
34 49 28.2 226 2 A87664 hypothetical prote  
35 49 28.2 772 2 T07958 protoporphyrin IX  
36 49 28.2 945 2 E83117 excinuclease ABC c  
37 49 28.2 952 1 T46550 excinuclease ABC c  
38 49 28.2 961 2 AE2270 excinuclease ABC c  
39 49 28.2 968 2 D87570 excinuclease ABC, c  
40 49 28.2 970 2 S77349 excinuclease ABC c  
41 49 28.2 1014 2 T36031 excinuclease ABC c  
42 49 28.2 1193 2 T50729 magnesium-protopor  
43 49 28.2 1328 2 AE2351 protoporphyrin IX  
44 49 28.2 1331 2 S75000 protoporphyrin IX  
45 49 28.2 1379 2 S37310 protoporphyrin IX

## ALIGNMENTS

### RESULT 1

JC7607

prolactin-releasing peptide - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 30-Jun-2001

C:Accession: JC7607

R:Yamada, M.; Ozawa, A.; Ishii, S.; Shibusawa, N.; Hashida, T.; Hosoya, T

Biochem. Biophys. Res. Commun. 281, 53-56, 2001

A:Title: Isolation and characterization of the rat prolactin-releasing peptide gene: Mu

A:Reference number: JC7607; MUID:21092785; PMID:11178959

A:Contents: Spleen

A:Accession: JC7607

A:Molecule type: DNA

A:Residues: 1-83 <YAM>

A:Cross-references: DDBJ:AB040612; DDBJ:AB040613

C:Comment: This peptide induces arachidonic acid metabolite release from rat anterior p

release, and stimulation of ACTH secretion from the pituitary.

C:Genetics:

A:Gene: PrRP

A:Introns: 33/1

Query Match 100.0%; Score 174; DB 2; Length 83;

Best Local Similarity 100.0%; Pred. No. 4.4e-18; Mismatches 0; Indels 0; Gaps 0;

Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRAHQSMETRTDINPAWYTGIRPVGRF 31

Db 22 SRAHQSMETRTDINPAWYTGIRPVGRF 52

### RESULT 2

T50904

Mg protoporphyrin methyl transferase [imported] - Rubrivivax gelatinosus

C:Species: Rubrivivax gelatinosus

C:Date: 21-Jul-2000 #sequence\_revision 21-Jul-2000 #text\_change 02-Sep-2000

C:Accession: T50904

R:Nagashima, K.V.; Igarashi, N.; Harada, J.; Nagashima, S.; Matsuura, K.; Shimada, K.

submitted to the EMBL Data Library, November 1999

A:Description: Determination of Nucleotide Sequences of Rubrivivax gelatinosus Photosyn

A:Reference number: Z25270

A:Accession: T50904

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1236 <NAG>

A:Cross-references: EMBL:AB034704; PIDN:BAA94057.1

A:Experimental source: strain Il144

C:Genetics:

C:Superfamily: Rhodobacter capsulatus magnesium-protoporphyrin O-methyltransferase

Query Match 35.1%; Score 61; DB 2; Length 1236;

Best Local Similarity 37.5%; Pred. No. 1.9;

```

Matches 12; Conservative 6; Mismatches 8; Indels 6; Gaps 1;

QY 3 AHQHSMETRPDINPAWYTG-----RGIRPV 28
      : : : : : : : : : : : : : : : : : :
Db 1112 SEQVALETRMLNPKWYEGMLHGYEGVRQI 1143

RESULT 3
T31462
probable magnesium chelatase (EC 4.99.1.-) chain H BchH - Helicobacillus mobilis
C/Species: Helicobacillus mobilis
C/Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 15-Sep-2000
C/Accession: T31462
R/Xiong, J.; Inoue, K.; Bauer, C.E.
Proc. Natl. Acad. Sci. U.S.A. 95, 14851-14856, 1998
A/Title: Tracking molecular evolution of photosynthesis by characterization of a major P
A/Reference number: Z21036; MUID:99061957; PMID:9843979
A/Accession: T31462
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1292 <XIO>
A/Cross-references: EMBL:AF080002; NID:g3820536; PID:g3820560; PIDN:AAC84033.1
C/Genetics:
A/Gene: bchH
C/Superfamily: Rhodobacter capsulatus magnesium-protoporphyrin O-methyltransferase
C/Keywords: lyase

Query Match 33.3%; Score 58; DB 2; Length 1292;
Best Local Similarity 42.3%; Pred. No. 5.5;
Matches 11; Conservative 3; Mismatches 6; Indels 6; Gaps 1;

QY 9 ETRTPDINPAWYTG-----RGIRPV 28
      : : : : : : : : : : : : : : : : : :
Db 1176 ETRKTLNPKWYEGMLKHGYEGVREI 1201

RESULT 4
H69157
excinuclease ABC chain A - Methanobacterium thermoautotrophicum (strain Delta H)
N/Alternate names: uvra protein
N/Contains: excision endonuclease ABC (EC 3.1.1.-) chain A
C/Species: Methanobacterium thermoautotrophicum
C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 02-Feb-2001
C/Accession: H69157
R/Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwanli, N.;
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A/Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
A/Reference number: A69000; MUID:98037514; PMID:9371463
A/Accession: H69157
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-962 <MTH>
A/Cross-references: GB:AE000828; GB:AE000666; NID:g2621504; PIDN:AAB84949.1; PID:g2621504
A/Experimental source: strain Delta H
C/Genetics:
A/Gene: MTH443
A/Start codon: TTG
C/Superfamily: excinuclease ABC chain A; ATP-binding cassette homology
C/Keywords: ATP; DNA binding; DNA repair; duplication; hydrolase; nucleotide binding; P-
F;38-45/Region: nucleotide-binding motif A (P-loop)
F:632-915/Domain: ATP-binding cassette homology <ABCE>
F:649-656/Region: nucleotide-binding motif A (P-loop)

Query Match 32.2%; Score 56; DB 2; Length 962;
Best Local Similarity 42.1%; Pred. No. 7.8;
Matches 16; Conservative 0; Mismatches 4; Indels 18; Gaps 2;

QY 11 RTPDINPAWYTG-----RGIREVGRF 31
      : : : : : : : : : : : : : : : : : :
Db 703 RTPRSNPATYGVTFTHIRELPAQTPARKGRYP-GRF 739

```



A:Cross-references: GB:BA000019; PIDN:BA078052.1; PID:gl17135506; GSPDB:GN00179  
 A:Experimental source: strain PCC 7120  
 C:Genetics:  
 A:Gene: all1686

Query Match 30.2%; Score 52.5; DB 2; Length 303;  
 Best Local Similarity 48.4%; Pred. No. 7.2;  
 Matches 15; Conservative 2; Mismatches 9; Indels 5; Gaps 2;

Qy 4 HQHSMETTPDINAWY-----TGIRPVGR 30  
 Db 226 HEHSYE-RTRAIIDGTYLTTCGAGAGRPVGR 255

RESULT 8  
 A82329  
 excinuclease ABC chain A [similarity] - Vibrio cholerae (strain N16961 serogroup O1)  
 N:Contains: excision endonuclease ABC (EC 3.1.-.-) chain A  
 C:Species: Vibrio cholerae  
 C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 24-Aug-2001  
 C:Accession: A82329  
 R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;  
 Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, B.  
 1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
 Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833; PMID:10952301  
 A:Accession: A82329  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-940 <HEI>

A:Cross-references: GB:AE004127; GB:AE003852; NID:G9654808; PIDN:AAF93567.1; GSPDB:GN001  
 A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
 C:Genetics:

A:Map position: 1  
 C:Superfamily: excinuclease ABC chain A; ATP-binding cassette homology  
 C:Keywords: ATP; DNA binding; DNA repair; hydrolase; nucleotide binding; P-loop  
 F:31-38/Region: nucleotide-binding motif A (P-loop)  
 F:639-646/Region: nucleotide-binding motif A (P-loop)

Query Match 29.9%; Score 52; DB 2; Length 940;  
 Best Local Similarity 39.5%; Pred. No. 29;  
 Matches 15; Conservative 1; Mismatches 4; Indels 18; Gaps 2;

Qy 11 RTPDINPAWYTG-----RGIRPVGRF 31  
 Db 694 RTPRSNPATYTGTFPIRELFAGTQESRSRGYP-GRF 730

RESULT 9  
 B87083  
 excinuclease ABC subunit A [imported] - Mycobacterium leprae  
 C:Species: Mycobacterium leprae  
 C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 10-May-2001  
 C:Accession: B87083  
 R:Coile, S.T.; Eigmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho  
 eam, M.A.; Rutherford, K.M.  
 Nature 409, 1007-1011, 2001  
 A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; S  
 A:Title: Massive gene decay in the leprosy bacillus.  
 A:Reference number: A86909; MUID:21128732; PMID:11234002  
 A:Accession: B87083  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-969 <STO>

A:Cross-references: GB:AL450380; NID:gl13093278; PIDN:CAC31773.1; GSPDB:GN00147  
 C:Genetics:  
 A:Gene: uvrA  
 C:Superfamily: excinuclease ABC chain A; ATP-binding cassette homology

Query Match 29.9%; Score 52; DB 2; Length 969;

Best Local Similarity 39.5%; Pred. No. 30;  
 Matches 15; Conservative 1; Mismatches 4; Indels 18; Gaps 2;  
 Qy 11 RTPDINPAWYTG-----RGIRPVGRF 31  
 Db 706 RTPRSNPATYTGVDKIRILFAATTEAKVRGYP-GRF 742

RESULT 10  
 A70619  
 excinuclease ABC chain A - Mycobacterium tuberculosis (strain H37RV)  
 N:Alternate names: uvrA protein  
 N:Contains: excision endonuclease ABC (EC 3.1.-.-) chain A  
 C:Species: Mycobacterium tuberculosis  
 C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 02-Feb-2001  
 C:Accession: A70619  
 R:Coile, S.T.; Broesch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S  
 ; Connor, R.M.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S  
 ; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
 Nature 393, 537-544, 1998  
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
 A:Reference number: A70500; MUID:98295987; PMID:9634230  
 A:Accession: A70619  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-972 <COL>

A:Cross-references: GB:Z85982; GB:AL123456; NID:G3261718; PIDN:CAB06633.1; PID:gl818989  
 A:Experimental source: strain H37RV  
 C:Genetics:  
 A:Gene: uvrA  
 C:Superfamily: excinuclease ABC chain A; ATP-binding cassette homology  
 C:Keywords: ATP; DNA binding; DNA repair; hydrolase; nucleotide binding; P-loop  
 F:32-39/Region: nucleotide-binding motif A (P-loop)  
 F:637-920/Domain: ATP-binding cassette homology <ABCE>  
 F:654-661/Region: nucleotide-binding motif A (P-loop)

Query Match 29.9%; Score 52; DB 2; Length 972;  
 Best Local Similarity 39.5%; Pred. No. 30;  
 Matches 15; Conservative 1; Mismatches 4; Indels 18; Gaps 2;

Qy 11 RTPDINPAWYTG-----RGIRPVGRF 31  
 Db 708 RTPRSNPATYTGVDKIRILFAATTEAKVRGYP-GRF 744

RESULT 11  
 AH3361  
 excinuclease ABC chain A BMEI0878 [imported] - Brucella melitensis (strain 16M)  
 C:Species: Brucella melitensis  
 C:Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 15-Feb-2002  
 C:Accession: AH3361  
 R:DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mijer, C.; Los, T.; Ivanova,  
 ; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letes  
 Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
 A:Title: The genome sequence of the facultative intracellular pathogen Brucella meliten  
 A:Reference number: AD3252; PMID:11756688  
 A:Accession: AH3361  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-974 <KUR>  
 A:Cross-references: GB:AE008917; PIDN:AAU52059.1; PID:gl7982827; GSPDB:GN00190  
 A:Experimental source: strain 16M  
 C:Genetics:  
 A:Gene: BMEI0878  
 A:Map position: 1  
 C:Superfamily: excinuclease ABC chain A; ATP-binding cassette homology

Query Match 29.9%; Score 52; DB 2; Length 974;  
 Best Local Similarity 39.5%; Pred. No. 30;  
 Matches 15; Conservative 1; Mismatches 4; Indels 18; Gaps 2;

Qy 11 RTPDINPAWYTG-----RGIRPVGRF 31



Matches 10; Conservative 4; Mismatches 15; Indels 3; Gaps 1;

Qy 1 SRAHQSMETRP---DINPAWYTGIRPVG 29

Db 120 SECHOHNFVYLPDAVDDLTQWFIAGFEQVG 151

Search completed: December 3, 2003, 19:14:56

Job time : 11.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 19:01:34 ; Search time 6.25 Seconds  
(without alignments)  
233.252 Million cell updates/sec

Title: US-09-868-885B-18

Perfect score: 174  
Sequence: 1 SRAHQHSMETRTDINPAWYGRGIRPVGRF 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID            | Description        |
|------------|-------|-------------|--------|---------------|--------------------|
| 1          | 174   | 100.0       | 83     | 1 PRRP_RAT    | P81278 rattus norv |
| 2          | 163   | 93.7        | 98     | 1 PRRP_BOVIN  | P81264 bos taurus  |
| 3          | 149   | 85.6        | 87     | 1 PRRP_HUMAN  | P81277 homo sapien |
| 4          | 56    | 32.2        | 428    | 1 NER3_BOVIN  | O97859 bos taurus  |
| 5          | 56    | 32.2        | 962    | 1 UVRA_METTH  | O26543 methanobact |
| 6          | 53    | 30.5        | 798    | 1 UNR_EAT     | P18395 rattus norv |
| 7          | 52    | 29.9        | 940    | 1 UVRA_VIBCH  | Q9Kuw5 vibrio chol |
| 8          | 52    | 29.9        | 969    | 1 UVRA_MYCLE  | Q8cc24 mycobacteri |
| 9          | 52    | 29.9        | 972    | 1 UVRA_MYCTU  | P94972 mycobacteri |
| 10         | 52    | 29.9        | 973    | 1 UVRA_RHILLO | Q98m36 rhizobium l |
| 11         | 52    | 29.9        | 973    | 1 UVRA_RHIME  | P56899 rhizobium m |
| 12         | 51    | 29.3        | 428    | 1 NER3_HUMAN  | Q9uq49 homo sapien |
| 13         | 51    | 29.3        | 943    | 1 UVRA_HAEIN  | P44410 haemophilus |
| 14         | 51    | 29.3        | 943    | 1 UVRA_PASMU  | P57979 pasteurella |
| 15         | 50.5  | 29.0        | 772    | 1 LMBT_HUMAN  | Q9y468 homo sapien |
| 16         | 50    | 28.7        | 264    | 1 X355_BUCAL  | P57436 buchnera ap |
| 17         | 50    | 28.7        | 482    | 1 R167_YEAST  | P39743 saccharomyc |
| 18         | 50    | 28.7        | 940    | 1 UVRA_ECO57  | Q8x5u9 escherichia |
| 19         | 50    | 28.7        | 940    | 1 UVRA_ECOL6  | Q8fb02 escherichia |
| 20         | 50    | 28.7        | 940    | 1 UVRA_ECOLI  | P07671 escherichia |
| 21         | 50    | 28.7        | 941    | 1 UVRA_SALTY  | P37434 salmonella  |
| 22         | 50    | 28.7        | 947    | 1 UVRA_YERPE  | Q8zj07 versinia pe |
| 23         | 50    | 28.7        | 953    | 1 UVRA_RICPR  | Q8zcc3 rickettsia  |
| 24         | 50    | 28.7        | 955    | 1 UVRA_RICCN  | Q8z931 rickettsia  |
| 25         | 49.5  | 28.4        | 374    | 1 YHHJ_ECOLI  | P31993 escherichia |
| 26         | 49    | 28.2        | 945    | 1 UVRA_PSEAE  | Q9hwg0 pseudomonas |
| 27         | 49    | 28.2        | 952    | 1 UVRA_THETH  | Q56242 thermus the |
| 28         | 49    | 28.2        | 970    | 1 UVRA_SYNV3  | P73412 synechocyst |
| 29         | 49    | 28.2        | 1014   | 1 UVRA_STRCO  | Q9z507 streptomyce |
| 30         | 49    | 28.2        | 1193   | 1 BCHH_RHOSH  | Q9rfd5 rhodobacter |
| 31         | 48.5  | 27.9        | 213    | 1 CYSC_YERPE  | Q8zbp3 versinia pe |
| 32         | 48    | 27.6        | 417    | 1 EX7L_CORGL  | Q8nmr3 corynebacte |
| 33         | 48    | 27.6        | 447    | 1 EX7L_COREF  | Q8fqpl corynebacte |

34 48 27.6 569 1 UVRA\_VITST Q08518 vitreoscill  
35 48 27.6 719 1 NRPI\_YEAST P32770 saccharomyc  
36 48 27.6 798 1 UNR\_HUMAN O75534 homo sapien  
37 48 27.6 948 1 UVRA\_NEIMA O9jue4 neisseria m  
38 48 27.6 949 1 UVRA\_NEIMB O9jzpl neisseria m  
39 48 27.6 950 1 UVRA\_NEIGO Q50968 neisseria g  
40 48 27.6 960 1 UVRA\_TREPA O83527 treponema p  
41 48 27.6 1087 1 XPO7\_HUMAN O9ui99 homo sapien  
42 48 27.6 1087 1 XPO7\_MOUSE Q9epk7 mus musculu  
43 48 27.6 1194 1 BCHH\_RHOCA P26162 rhodobacter  
44 47.5 27.3 501 1 TRA2\_MOUSE P39429 mus musculu  
45 47.5 27.3 941 1 GCSP\_MYCTU Q50601 mycobacteri

#### ALIGNMENTS

RESULT 1  
PRRP\_RAT  
ID PRRP\_RAT STANDARD; PRT; 83 AA.  
AC P81278; Q8K3Y0;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DB Prolactin-releasing peptide precursor (PrRP) (Prolactin-releasing hormone) [Contains: Prolactin-releasing peptide PrRP31; Prolactin-releasing peptide PrRP20].  
GN PRH.  
OS Rattus norvegicus (Rat).  
OC Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RC TISSUE=Brain;  
RX MEDLINE=98268781; PubMed=9607765;  
RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S., Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M., Kurokawa T., Nishimura O., Onda H., Fujino M.;  
RA "A prolactin-releasing peptide in the brain.";  
RT Nature 393:272-276(1998).  
RL [2]  
RN SEQUENCE FROM N.A. (ISOFORM 2).  
RP STRAIN=Sprague-Dawley; TISSUE=Hypothalamus;  
RC Anderson S.T., Kokay I.C., Lang T., Grattan D.R., Curlewis J.D.;  
RT "Quantitation of prolactin-releasing peptide (PrRP) mRNA expression in specific brain regions during the rat oestrous cycle and in lactation.";  
RT Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP TISSUE SPECIFICITY.  
RX MEDLINE=99426652; PubMed=10498338;  
RA Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Habata Y., Hinuma S., Sekiguchi M., Kitada C., Kurokawa T., Nishimura O., Onda H., Sumino Y., Fujino M.;  
RA "Tissue distribution of prolactin-releasing peptide (PrRP) and its receptor.";  
RT Regul. Pept. 83:1-10(1999).  
RL [4]  
CC -1- FUNCTION: Stimulates prolactin (PRL) release and regulates the expression of prolactin through its receptor GPR10. May stimulate lactotrophs directly to secrete PRL.  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event-Alternative splicing; Named isoforms=2;  
CC Name=1;  
CC IsoId=P81278-1; Sequence=Displayed;  
CC Name=2;  
CC IsoId=P81278-2; Sequence=VSP\_004370;  
CC -1- TISSUE SPECIFICITY: Widely expressed, with highest levels in medulla oblongata and hypothalamus.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its



Qy 2 RAHQSMETRTPDINPAWYTGIRPV 28  
||| : : : : :  
Db 195 RARPHSLMIYSDDLGATWHGRLIKPM 221

```

RESULT 5
ID UVRA METHTH STANDARD; PRT; 962 AA.
AC 026543;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE UvrABC system protein A (UvrA protein) (Excinuclease ABC subunit A).
GN UVRA OR MTH443.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;
NN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Delta H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jivani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT *Complete genome sequence of Methanobacterium thermoautotrophicum
RT deitah: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
CC -1- FUNCTION: The UvrABC repair system catalyzes the recognition and
CC processing of DNA lesions. UvrA is an ATPase and a DNA-binding
CC protein. A damage recognition complex composed of 2 uvrA and 2
CC uvrB subunits scans DNA for abnormalities. When the presence of a
CC lesion has been verified by uvrB, the uvrA molecules dissociate
CC (by similarity).
CC -1- SUBUNIT: Forms a heterotetramer with uvrB during the search for
CC lesions (by similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. UVRA SUBFAMILY
-----
This SWISS-PROT entry is copyright. It is produced through a collabora-
tion between the Swiss Institute of Bioinformatics and the EMBL outstation
in the European Bioinformatics Institute. There are no restrictions on
use by non-profit institutions as long as its content is in
modified and this statement is not removed. Usage by and for commu-
nities requires a license agreement (See http://www.ebi.ac.uk/submit/
or send an email to license@ebi.ac.uk).
-----
EMBL: AE000828; AAB84949.1; -
DR F1R; H69157; H69157.
DR HAMAP; MF_00205; -; 1.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR003439; ABC transporter.
DR InterPro; IPR004602; UvrA.
DR Pfam; PF00005; ABC tran; 2.
DR ProDom; PD000006; ABC transporter; 1.
DR SMART; SMO0382; AAA; 2.
DR TIGRFAMs; TIGR00630; uvrA; 1.
DR PROSITE; PS00211; ABC TRANSPORTER_1; 2.
DR PROSITE; PS00893; ABC TRANSPORTER_2; 2.
KW SOS response; Excision nuclease; DNA repair; DNA recombination;
KW DNA excision; ATP-binding; DNA-binding; Repeat; Zinc; Metal-binding;
KW Zinc-finger; Complete proteome.
FT NP BIND 38 45 ATP (POTENTIAL) .
FT NP BIND 649 656 ATP (POTENTIAL) .
FT ZN FING 748 774 CA-TYPE.
FT SEQUENCE 962 AA; 108395 MW; 2C0E7FC41CCD060 CRC64;
Query Match 32.2%; Score 56; DB 1; Length 962;
Best Local Similarity 42.1%; Pred. No. 2.6;

```

Query Match 32.2%; Score 56; DB 1; Length 428;  
Best Local Similarity 37.0%; Pred. No. 1.2;  
Matches 10; Conservative 6; Mismatches 11; Indels 0; Gaps 0;



AC Q9CC24;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE UvrABC system protein A (UvrA protein) (Excinuclease ABC subunit A).  
 GN UVR A OR ML1392.  
 OS Mycobacterium leprae.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1769;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TN;  
 RX MEDLINE=21128732; PubMed=11234002;  
 RA Cole S.T., Eigmeier K., Parkhill J., James K.D., Thomson N.R.,  
 RA Wheeler P.R., Honore D., Brown D., Chillingworth T., Connor R.,  
 RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,  
 RA Holroyd S., Hornby T., Jagels K., Letwolt C., Maclean J., Moule S.,  
 RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,  
 RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,  
 RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,  
 RA Barrell B.G.;  
 RT "Massive gene decay in the leprosy bacillus."  
 RL Nature 409:1007-1011(2001).  
 CC -!- FUNCTION: The UvrABC repair system catalyzes the recognition and  
 CC processing of DNA lesions. UvrA is an ATPase and a DNA-binding  
 CC protein. A damage recognition complex composed of 2 uvrA and 2  
 CC uvrB subunits scans DNA for abnormalities. When the presence of a  
 CC lesion has been verified by uvrB, the uvrA molecules dissociate  
 CC (By similarity).  
 CC -!- SUBUNIT: Forms a heterotetramer with uvrB during the search for  
 CC lesions (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. UVR A SUBFAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AL583921; CAC31773.1; -.  
 DR PIR; B87083; B87083.  
 DR Leproma; ML1392; -.  
 DR HAMAP; MF\_00205; -; 1.  
 DR InterPro; IPR003439; ABC transporter.  
 DR InterPro; IPR004602; UvrA.  
 DR Pfam; PF00005; ABC tran; 2.  
 DR TIGRFAMs; TIGR00630; uvrA; 1.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER\_1; 2.  
 DR PROSITE; PS50893; ABC\_TRANSPORTER\_2; 1.  
 KW SOS response; Excision nuclease; DNA repair; DNA recombination;  
 KW DNA excision; ATP-binding; DNA-binding; Repeat; Zinc; Metal-binding;  
 KW Zinc-finger; Complete proteome.  
 FT NP\_BIND 32 39 ATP (POTENTIAL).  
 FT ZN\_FING 258 286 C4-TYPE.  
 FT NP\_BIND 522 659 ATP (POTENTIAL).  
 FT ZN\_FING 751 777 C4-TYPE.  
 SQ SEQUENCE 969 AA; 106377 MW; 6668141833C53DB5 CRC64;  
 Query Match 29.9%; Score 52; DB 1; Length 969;  
 Best Local Similarity 39.5%; Pred. No. 11;  
 Matches 15; Conservative 1; Mismatches 4; Indels 18; Gaps 2;  
 QY 11 RTPDINPAWYTC-----RGIRPVGRF 31  
 DB 706 RTPSPNPATYTGVDKIRILFAATTEAKVRGVP-QRF 742

RESULT 9  
 UVR A MYCTU  
 ID UVR A MYCTU STANDARD; PRT; 972 AA.  
 AC P94972;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE UvrABC system protein A (UvrA protein) (Excinuclease ABC subunit A).  
 GN UVR A OR RV1638 OR MT1675 OR MTCY06H11.02.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1773;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=H37Rv;  
 RX MEDLINE=98295987; PubMed=9634230;  
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,  
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R., Davlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,  
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers R.,  
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
 RA Sultston J.E., Taylor K., Whitehead S., Barrell B.G.;  
 RT "Deciphering the biology of Mycobacterium tuberculosis from the  
 RT complete genome sequence."  
 RL Nature 393:537-544(1998).  
 CC [2]  
 CC SEQUENCE FROM N.A.  
 RC STRAIN=CDC 1551 / Oshkosh;  
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
 RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,  
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,  
 RA Decher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,  
 RA Bishai W.;  
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and  
 RT laboratory strains."  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: The UvrABC repair system catalyzes the recognition and  
 CC processing of DNA lesions. UvrA is an ATPase and a DNA-binding  
 CC protein. A damage recognition complex composed of 2 uvrA and 2  
 CC uvrB subunits scans DNA for abnormalities. When the presence of a  
 CC lesion has been verified by uvrB, the uvrA molecules dissociate  
 CC (By similarity).  
 CC -!- SUBUNIT: Forms a heterotetramer with uvrB during the search for  
 CC lesions (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. UVR A SUBFAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; Z85982; CAB06633.1; -.  
 DR EMBL; AB007031; AAK45944.1; ALT\_INIT.  
 DR PIR; A70619; A70619.  
 DR TIGR; MT1675; -.  
 DR Tuberculist; RV1638; -.  
 DR HAMAP; MF\_00205; -; 1.  
 DR InterPro; IPR003439; ABC transporter.  
 DR InterPro; IPR004602; UvrA.  
 DR Pfam; PF00005; ABC tran; 2.  
 DR ProDom; PD000006; ABC transporter; 1.  
 DR TIGRFAMs; TIGR00630; uvrA; 1.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER\_1; 2.  
 DR PROSITE; PS50893; ABC\_TRANSPORTER\_2; 1.  
 KW SOS response; Excision nuclease; DNA repair; DNA recombination;  
 KW DNA excision; ATP-binding; DNA-binding; Repeat; Zinc; Metal-binding;  
 KW DNA excision; ATP-binding; DNA-binding; Repeat; Zinc; Metal-binding;



```
KW Zinc-finger; Complete proteome.
FT NP_BIND 32 39 ATP (POTENTIAL).
FT ZN_FING 257 285 C4-TYPE (ATYPICAL).
FT NP_BIND 654 661 ATP (POTENTIAL).
FT ZN_FING 753 779 C4-TYPE.
SQ SEQUENCE 972 AA; 106131 MW; 8937A764E592D981 CRC64;

Query Match 29.9%; Score 52; DB 1; Length 972;
Best Local Similarity 39.5%; Pred. No. 11;
Matches 15; Conservative 1; Mismatches 4; Indels 18; Gaps 2;

QY 11 RTPDINPAWYTG-----RGIRPVGRF 31
||| ||| |||
DB 708 RTPRNPATYTGVDKIRTLPAATTEAKVRGYP-CRF 744
||| ||| |||

RESULT 10
UVRA_RHILO UVRA_RHILO STANDARD; PRT; 973 AA.
AC Q98M36;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE UvrABC system protein A (UvrA protein) (Excinuclease ABC subunit A).
GN UVRA OR MLR0750.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Iidesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti";
RL DNA Res. 7:331-338(2000).
CC -1- FUNCTION: The UvrABC repair system catalyzes the recognition and
CC processing of DNA lesions. UvrA is an ATPase and a DNA-binding
CC protein. A damage recognition complex composed of 2 uvrA and 2
CC uvrB subunits scans DNA for abnormalities. When the presence of a
CC lesion has been verified by uvrB, the uvrA molecules dissociate
CC (By similarity).
CC -1- SUBUNIT: Forms a heterotetramer with uvrB during the search for
CC lesions (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. UVRA SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AP002995; BAB48277.1; -.
CC HAMAP; MF 00205; -.
CC InterPro; IPR003439; ABC transporter.
CC InterPro; IPR004602; UvrA.
CC Pfam; PF00005; ABC tran; 2.
CC TIGRfams; TIGR00630; uvrA; 1.
CC PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
CC PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
KW SOS response; Excision nuclease; DNA repair; DNA recombination;
KW DNA excision; ATP-binding; DNA-binding; Repeat; Zinc; Metal-binding;
KW Zinc-finger; Complete proteome.
FT NP_BIND 34 41 ATP (POTENTIAL).
FT ZN_FING 257 285 C4-TYPE (ATYPICAL).
FT NP_BIND 654 661 ATP (POTENTIAL).
FT ZN_FING 753 779 C4-TYPE.
SQ SEQUENCE 972 AA; 106131 MW; 8937A764E592D981 CRC64;

Query Match 29.9%; Score 52; DB 1; Length 972;
Best Local Similarity 39.5%; Pred. No. 11;
Matches 15; Conservative 1; Mismatches 4; Indels 18; Gaps 2;

QY 11 RTPDINPAWYTG-----RGIRPVGRF 31
||| ||| |||
DB 708 RTPRNPATYTGVDKIRTLPAATTEAKVRGYP-CRF 744
||| ||| |||

RESULT 11
UVRA_RHIME UVRA_RHIME STANDARD; PRT; 973 AA.
AC P56899;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE UvrABC system protein A (UvrA protein) (Excinuclease ABC subunit A).
GN UVRA OR R01557 OR SMC01235.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396507; PubMed=11481430;
RA Capella D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe P., Batut J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kles B., Lelaure V., Masuy D.,
RA Pohl T., Portetle D., Puehler A., Purnelle B., Ramsperger U.,
RA Renard C., Thebaud P., Vandenbol M., Weidner S., Gallibert P.;
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
RN [2]
RP SEQUENCE OF 1-140 FROM N.A.
RC STRAIN=2021;
RX MEDLINE=99430868; PubMed=10503543;
RA Tapas A., Barbe J.;
RT "Regulation of divergent transcription from the uvrA-sb promoters in
RT Sinorhizobium meliloti.";
RL Mol. Gen. Genet. 262:121-130(1999).
CC -1- FUNCTION: The UvrABC repair system catalyzes the recognition and
CC processing of DNA lesions. UvrA is an ATPase and a DNA-binding
CC protein. A damage recognition complex composed of 2 uvrA and 2
CC uvrB subunits scans DNA for abnormalities. When the presence of a
CC lesion has been verified by uvrB, the uvrA molecules dissociate
CC (By similarity).
CC -1- SUBUNIT: Forms a heterotetramer with uvrB during the search for
CC lesions (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. UVRA SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AL591787; CAC46136.1; -.
CC EMBL; AF125162; AAF03210.1; -.
CC HAMAP; MF 00205; -.
CC InterPro; IPR003439; ABC transporter.
CC InterPro; IPR004602; UvrA.
CC Pfam; PF00005; ABC tran; 2.
CC ProDom; PD000006; ABC transporter; 1.
CC TIGRfams; TIGR00630; uvrA; 1.
```

[illegible]

```

RP SEQUENCE OF 1-71 FROM N.A.
RC STRAIN=NTHI TN106;
RX MEDLINE=94341556; PubMed=8063092;
RA Jarosik G.P., Hansen E.J.;
RT "Cloning and sequencing of the Haemophilus influenzae ssb gene
RL encoding single-strand DNA-binding protein."
Gene 146:101-103(1994).
CC -1- FUNCTION: The UvrABC repair system catalyzes the recognition and
CC processing of DNA lesions. UvrA is an ATPase and a DNA-binding
CC protein. A damage recognition complex composed of 2 uvrA and 2
CC uvrB subunits scans DNA for abnormalities. When the presence of a
CC lesion has been verified by uvrB, the uvrA molecules dissociate
CC (By similarity).
CC -1- SUBUNIT: Forms a heterotrimer with uvrB during the search for
CC lesions (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. UVRB SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; U32711; AAC21915.1; -
DR EMBL; U33877; AAC44592.1; -
DR EMBL; U04997; AAA60462.1; -
DR FIC; D64057; D64057.
DR TIGR; H10249; -.
DR HAMAP; MF 00205; -; 1.
DR InterPro; IPR003439; ABC transporter.
DR InterPro; IPR004602; UvrA.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 1.
DR TIGRFAMs; TIGR00630; uvrA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
KW SOS response; Excision nuclease; DNA repair; DNA recombination;
KW DNA excision; ATP-binding; DNA-binding; Repeat; Zinc; Metal-binding;
KW Zinc-finger; Complete proteome.
FT NP_BIND 31 38 ATP (POTENTIAL).
FT ZN_FING 253 280 C4-TYPE.
FT NP_BIND 640 647 ATP (POTENTIAL).
FT ZN_FING 740 766 C4-TYPE.
FT ZN_FING 163 163 V -> L (IN REF. 2).
FT CONFLICT 236 236 E -> D (IN REF. 2).
FT CONFLICT 425 425 R -> K (IN REF. 2).
FT CONFLICT 463 463 I -> M (IN REF. 2).
FT CONFLICT 514 514 E -> Q (IN REF. 2).
FT CONFLICT 661 661 A -> T (IN REF. 2).
FT CONFLICT 928 928 T -> E (IN REF. 2).
FT CONFLICT 935 942 FLKPLEK -> EL (IN REF. 2).
SQ SEQUENCE 943 AA; 104366 MW; 4DBA0DCA602D465 CRC64;

Query Match 29.3%; Score 51; DB 1; Length 943;
Best Local Similarity 39.5%; Pred. No. 15;
Matches 15; Conservative 0; Mismatches 5; Indels 18; Gaps 2;

QY 11 RTPDINPAWYTG-----RGIRPVGRF 31
Db 695 RTRSPNPATYTGFTPIRELFAGVPESRARGYNP-GRF 731

RESULT 14
UVRB_PASMU
ID UVRB_PASMU STANDARD; PRT; 943 AA.
AC P57979;
DT 16-OCT-2001 (Rel. 40, Created)
DT 18-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE UvrABC system protein A (UvrA protein) (Excinuclease ABC subunit A).

```

```

GN UVRB OR PM1951.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pm70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Faustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70."
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -1- FUNCTION: The UvrABC repair system catalyzes the recognition and
CC processing of DNA lesions. UvrA is an ATPase and a DNA-binding
CC protein. A damage recognition complex composed of 2 uvrA and 2
CC uvrB subunits scans DNA for abnormalities. When the presence of a
CC lesion has been verified by uvrB, the uvrA molecules dissociate
CC (By similarity).
CC -1- SUBUNIT: Forms a heterotrimer with uvrB during the search for
CC lesions (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. UVRB SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AF006231; AAK04035.1; -
DR HAMAP; MF 00205; -; 1.
DR InterPro; IPR003439; ABC transporter.
DR InterPro; IPR004602; UvrA.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 1.
DR TIGRFAMs; TIGR00630; uvrA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
KW SOS response; Excision nuclease; DNA repair; DNA recombination;
KW DNA excision; ATP-binding; DNA-binding; Repeat; Zinc; Metal-binding;
KW Zinc-finger; Complete proteome.
FT NP_BIND 31 38 ATP (POTENTIAL).
FT ZN_FING 253 280 C4-TYPE.
FT NP_BIND 640 647 ATP (POTENTIAL).
FT ZN_FING 740 766 C4-TYPE.
SQ SEQUENCE 943 AA; 104186 MW; A096DB1162D3C354 CRC64;

Query Match 29.3%; Score 51; DB 1; Length 943;
Best Local Similarity 39.5%; Pred. No. 15;
Matches 15; Conservative 0; Mismatches 5; Indels 18; Gaps 2;

QY 11 RTPDINPAWYTG-----RGIRPVGRF 31
Db 695 RTRSPNPATYTGFTPIRELFAGVPESRARGYNP-GRF 731

RESULT 15
LMBT_HUMAN
ID LMBT_HUMAN STANDARD; PRT; 772 AA.
AC Q9Y468; Q9H1B6; Q9H1G5; Q9UG06; Q9UJB9; Q9Y4C9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Lethal(3)malignant brain tumor-like protein (L(3)mbt-like) (L(3)mbt
DE protein homolog) (H-1(3)mbt protein) (H-L(3)MBT).
GN L3MBTL OR L3MBT OR KIAA0681.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

```

RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
RC TISSUE=Brain;  
RX MEDLINE=9373015; PubMed=1045843;  
RA Koga H., Matsui S.-I., Hirota T., Takebayashi S.-I., Okumura K.,  
RA Sava H.;  
RT "A human homolog of Drosophila lethal(3)malignant brain tumor  
RT (l3)mbt protein associates with condensed mitotic chromosomes.";  
RL Oncogene 18:3799-3809(1999).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORM 3).  
RC TISSUE=Uterus;  
RA Koehrer K., Beyer A., Mewes H.-W., Gassenhuber J., Wiemann S.;  
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21638749; PubMed=11780052;  
RA Deloukas P., Matthews L.H., Ashurat J., Burton J., Gilbert J.G.R.,  
RA Jones M., Stavridis G., Almeida J.P., Babbage A.K., Bagguley C.L.,  
RA Bailey J., Barlow K.P., Bates K.N., Beard L.M., Beare D.M.,  
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgman A.M., Brown A.J.,  
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
RA Chapman J.C., Clamp M., Clark A.P., Clark L.N., Clark S.Y., Clee C.M.,  
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,  
RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,  
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
RA Grahame D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
RA Kay M.F., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
RA Lehaevaiah M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
RA Mareh V.L., Martin S.L., McConnachie L.J., McIlroy K., McMurray A.A.,  
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
RA Phillips B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,  
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkhen R., Sims S.,  
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,  
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M., Williams S.A.,  
RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,  
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
RA Rogers J.;  
RT "The DNA sequence and comparative analysis of human chromosome 20.";  
RL Nature 414:865-871(2001).  
RN [4]  
RP SEQUENCE OF 215-772 FROM N.A. (ISOFORM 4).  
RC TISSUE=Brain;  
RX MEDLINE=98403880; PubMed=9734811;  
RA Ishikawa K.-I., Nagase T., Suyama M., Miyajima N., Tanaka A.,  
RA Kotani H., Nomura N., Ohara O.;  
RT "Prediction of the coding sequences of unidentified human genes. X.  
RT The complete sequences of 100 new cDNA clones from brain which can  
RT code for large proteins in vitro.";  
RL DNA Res. 5:169-176(1998).  
CC -1- FUNCTION: Polycarb group (PCG) protein. PCG proteins maintain the  
CC transcriptionally repressive state of genes, probably via a  
CC modification of chromatin, rendering it heritably changed in its  
CC expressibility. Probably plays a role in cell proliferation.  
CC Overexpression induces multinucleated cells, suggesting that it is  
CC required to accomplish normal mitosis.  
CC -1- SUBCELLULAR LOCATION: Nuclear; excluded from the nucleolus. Does  
CC not colocalizes with the PCG protein BML1, suggesting that these  
CC two proteins do not belong to the same complex.  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=4;  
CC Name=1; Synonyms=mbt-1;  
CC IsoId=Q3Y468-1; Sequence=Displayed;  
CC Name=2; Synonyms=mbt-1f;  
CC IsoId=Q3Y468-2; Sequence=VSP\_003902;  
CC Name=3;  
CC IsoId=Q3Y468-3; Sequence=VSP\_003901, VSP\_003902;  
CC Name=4;  
CC IsoId=Q3Y468-4; Sequence=VSP\_003903;  
CC -1- TISSUE SPECIFICITY: Widely expressed. Expression is reduced in

colorectal cancer cell line SW480 and promyelocytic leukemia cell  
line HL-60.  
-1- DEVELOPMENTAL STAGE: In interphase cells, it is scattered  
throughout the nucleoplasm. In mitotic cells, it strongly  
associates with condensed chromosomes from the prophase to  
telophase.  
-1- SIMILARITY: Contains 3 mbt domains.  
-1- CAUTION: Ref.3 (CAC18508) sequences differ from that shown due to  
erroneous gene model prediction.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; U89358; AAC69438.1; --  
EMBL; AL110279; CAB53714.1; --  
EMBL; Z98752; CAC16799.1; --  
EMBL; Z98752; CAC16800.1; --  
EMBL; Z98752; CAC18508.1; --  
EMBL; AL031681; CAC17518.1; ALT\_SEQ.  
EMBL; AB014581; CAB43959.1; --  
EMBL; AB014581; BAA31656.1; --  
PIR; T14794; T14794.  
Genew; HGNC:15905; L3MBTL.  
DR EMBL; U89358; AAC69438.1; --  
DR EMBL; AL110279; CAB53714.1; --  
DR EMBL; Z98752; CAC16799.1; --  
DR EMBL; Z98752; CAC16800.1; --  
DR EMBL; Z98752; CAC18508.1; --  
DR EMBL; AL031681; CAC17518.1; ALT\_SEQ.  
DR EMBL; AB014581; CAB43959.1; --  
DR EMBL; AB014581; BAA31656.1; --  
DR PIR; T14794; T14794.  
DR Genew; HGNC:15905; L3MBTL.  
DR GK; Q9Y468; --  
DR InterPro; IPR004092; MBT.  
DR InterPro; IPR002515; Znf\_C2HC.  
DR Pfam; PF02820; MBT1\_3.  
DR Pfam; PF01530; zf-C2HC; 1.  
DR SMART; SM00561; MBT; 3.  
DR Transcription regulation; Chromatin regulator; Zinc-finger;  
KW DNA-binding; Nuclear protein; Repeat; Alternative splicing.  
FT DOMAIN 242 315 MBT 1.  
FT DOMAIN 349 422 MBT 2.  
FT DOMAIN 453 526 MBT 3.  
FT ZN FING 552 578 C2HC-TYPE.  
FT VARSPLIC 1 348 Missing (in isoform 3).  
FT VARSPLIC 709 772 /FTId=VSP\_003901.  
FT ARIVRVTHVSGKTLVMTVAQLGDLVCSDLQBGKILETVG  
FT HSLCCLPHTLLAKLSPASDSQY -> VRCKRVDRAVTV  
FT VKTAGRCPPQRHFC (in isoform 2 and  
FT isoform 3).  
FT VARSPLIC 709 772 /FTId=VSP\_003902.  
FT ARIVRVTHVSGKTLVMTVAQLGDLVCSDLQBGKILETVG  
FT HSLCCLPHTLLAKLSPASDSQY -> MIDGEAPLLLTQAD  
FT IVKIMSVKLGPAKLIYNAILMFKNADDTLK (in  
FT isoform 4).  
FT /FTId=VSP\_003903.  
FT P -> L (IN REF. 1).  
FT LR -> MC (IN REF. 1).  
FT L -> M (IN REF. 1).  
FT S -> P (IN REF. 1).  
SQ SEQUENCE 772 AA; 85916 MW; 117B03A628826B29 CRC64;  
Query Match 29.0%; Score 50.5; DB 1; Length 772;  
Best Local Similarity 36.7%; Pred. No. 14;  
Matches 11; Conservative 6; Mismatches 10; Indels 3; Gaps 2;  
OY 1 SRAHQHSMETPTDINPA-WY--TGGRIRP 27  
Db 276 SECHDFWVNANSPDHPAGNPEKTKHKLQP 305  
Search completed: December 3, 2003, 19:12:02  
Job time : 6.25 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 19:03:25 ; Search time 26.5 Seconds  
(without alignments)  
301.873 Million cell updates/sec

Title: US-09-868-885B-18

Perfect score: 174

Sequence: 1 SRAHQSMETRTDINPAWYTGIRPVGRF 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: sp archaea.\*

2: sp bacteria.\*

3: sp fungi.\*

4: sp human.\*

5: sp invertebrate.\*

6: sp mammal.\*

7: sp mbc.\*

8: sp organelle.\*

9: sp phage.\*

10: sp plant.\*

11: sp rodent.\*

12: sp virus.\*

13: sp vertebrate.\*

14: sp unclassified.\*

15: sp\_rvirus.\*

16: sp\_bacteriap.\*

17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID     | Description        |
|------------|-------|-------------|--------|-----------|--------------------|
| 1          | 163   | 93.7        | 98     | 6 Q8WN12  | Q8WN12 oviv aries  |
| 2          | 107   | 61.5        | 117    | 13 Q9W624 | Q9W624 carassius a |
| 3          | 68    | 39.1        | 692    | 2 Q9L8U6  | Q9L8U6 rhodospiril |
| 4          | 61    | 35.1        | 1236   | 2 Q9TPA4  | Q9TPA4 rhodocyclu  |
| 5          | 58    | 33.3        | 1253   | 2 Q8TV77  | Q8TV77 uncultured  |
| 6          | 58    | 33.3        | 1292   | 2 Q9ZG85  | Q9ZG85 heliobacill |
| 7          | 56    | 32.2        | 355    | 5 Q9SYJ8  | Q9SYJ8 ciona savig |
| 8          | 53    | 30.5        | 54     | 4 Q9UJF9  | Q9UJF9 homo sapien |
| 9          | 53    | 30.5        | 139    | 11 Q8BHP9 | Q8BHP9 mus musculu |
| 10         | 53    | 30.5        | 223    | 2 Q85605  | Q85605 prevotella  |
| 11         | 53    | 30.5        | 465    | 4 Q80687  | Q80687 homo sapien |
| 12         | 53    | 30.5        | 465    | 4 Q8W855  | Q8W855 homo sapien |
| 13         | 53    | 30.5        | 689    | 11 Q8R3R1 | Q8R3R1 mus musculu |
| 14         | 53    | 30.5        | 767    | 11 Q8JZN2 | Q8JZN2 mus musculu |
| 15         | 53    | 30.5        | 798    | 11 Q91W50 | Q91W50 mus musculu |
| 16         | 53    | 30.5        | 927    | 16 Q8R8M5 | Q8R8M5 thermoanaer |

|    |      |      |      |           |                    |
|----|------|------|------|-----------|--------------------|
| 17 | 53   | 30.5 | 939  | 16 Q8XN15 | Q8XN15 clostridium |
| 18 | 53   | 30.5 | 940  | 16 Q8DCJ3 | Q8DCJ3 vibrio vuln |
| 19 | 53   | 30.5 | 949  | 16 Q8NQO0 | Q8NQO0 corynebacte |
| 20 | 53   | 30.5 | 954  | 16 Q8FTQ6 | Q8FTQ6 corynebacte |
| 21 | 53   | 30.5 | 1245 | 2 Q8K257  | Q8K257 uncultured  |
| 22 | 53   | 30.5 | 1415 | 16 Q9HVI8 | Q9HVI8 pseudomonas |
| 23 | 53   | 30.5 | 1845 | 17 Q8TTS7 | Q8TTS7 methanosarc |
| 24 | 52.5 | 30.2 | 303  | 16 Q8YWC7 | Q8YWC7 anabaena sp |
| 25 | 52.5 | 30.2 | 416  | 16 Q8K746 | Q8K746 streptococc |
| 26 | 52   | 29.9 | 503  | 10 Q9LSC6 | Q9LSC6 arabidopsis |
| 27 | 52   | 29.9 | 974  | 16 Q8YHC4 | Q8YHC4 brucella me |
| 28 | 52   | 29.9 | 974  | 16 Q8G0I9 | Q8G0I9 brucella su |
| 29 | 52   | 29.9 | 1326 | 16 Q8DM52 | Q8DM52 synchococc  |
| 30 | 51.5 | 29.6 | 503  | 16 Q9KXY1 | Q9KXY1 vibrio chol |
| 31 | 51   | 29.3 | 294  | 5 Q19530  | Q19530 caenorhabdi |
| 32 | 51   | 29.3 | 333  | 4 Q96SD4  | Q96SD4 homo sapien |
| 33 | 51   | 29.3 | 367  | 4 Q8NGQ2  | Q8NGQ2 homo sapien |
| 34 | 51   | 29.3 | 367  | 4 Q8IXT2  | Q8IXT2 homo sapien |
| 35 | 51   | 29.3 | 370  | 11 Q8CGW9 | Q8CGW9 mus musculu |
| 36 | 51   | 29.3 | 464  | 5 Q8SS78  | Q8SS78 encephalito |
| 37 | 50.5 | 29.0 | 176  | 3 Q88689  | Q88689 saccharomyc |
| 38 | 50.5 | 29.0 | 333  | 16 Q9RJI0 | Q9RJI0 streptomyce |
| 39 | 50.5 | 29.0 | 380  | 10 Q8L8A7 | Q8L8A7 arabidopsis |
| 40 | 50.5 | 29.0 | 414  | 2 Q33480  | Q33480 propionibac |
| 41 | 50.5 | 29.0 | 461  | 16 Q8NNC6 | Q8NNC6 corynebacte |
| 42 | 50.5 | 29.0 | 548  | 10 Q9LFA0 | Q9LFA0 arabidopsis |
| 43 | 50.5 | 29.0 | 772  | 4 Q8IUW7  | Q8IUW7 homo sapien |
| 44 | 50   | 28.7 | 514  | 5 Q9VRV3  | Q9VRV3 drosophila  |
| 45 | 50   | 28.7 | 940  | 16 Q8FB02 | Q8FB02 escherichia |

#### ALIGNMENTS

#### RESULT 1

Q8WN12 PRELIMINARY; PRT; 98 AA.  
AC Q8WN12  
DT 01-MAR-2002 (TREMBLrel. 20, Created)  
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DB Preprolactin-releasing peptide.  
OS Ovis aries (Sheep).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Ruminantia; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Caprinae; Ovis.  
OX NCBI\_TaxID=9940;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Curlewis J.D., Kusters D.H.L., Barclay J.L., Anderson S.T.;  
RT "Prolactin-releasing peptide (PRP) in the ewe: cDNA cloning, mRNA  
distribution and effects on prolactin secretion in vitro and in  
vivo."  
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF450453; AAL47178.1; -  
SQ SEQUENCE 98 AA; 10513 MW; 2A533331E62CA85 CRC64;

Query Match 93.7%; Score 163; DB 6; Length 98;  
Best Local Similarity 93.5%; Pred. No. 5.7e-16;  
Matches 29; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Cy 1 SRAHQSMETRTDINPAWYTGIRPVGRF 31  
|||||  
Db 23 SRAHQSMETRTDINPAWYTGIRPVGRF 53  
|||||

#### RESULT 2

Q9W624 PRELIMINARY; PRT; 117 AA.  
ID Q9W624  
AC Q9W624  
DT 01-NOV-1999 (TREMBLrel. 12, Created)  
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)



DR InterPro; IPR003672; COBN/Mg chltase.

DR Pfam; PF02514; COBN-Mg chel.; 1;  
SQ SEQUENCE 1253 AA; 137586 MW; C98ABC010E261511 CRC64;

Query Match 33.3%; Score 58; DB 2; Length 1253;

Best Local Similarity 37.5%; Pred. No. 15;

Matches 12; Conservative 5; Mismatches 9; Indels 6; Gaps 1;

QY 3 AHQSHMETRTPDINPAWYTG-----RGIRPV 28

DB 1129 AEQVALETRVLPNPKWYESMLDHGVEGVRAI 1160

RESULT 6

Q9ZG55

ID Q9ZG55 PRELIMINARY; PRT; 1292 AA.

AC Q9ZG55

DT 01-MAY-1999 (TRENBLrel. 10, Created)

DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)

DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)

DE Mg chelate subunit H BchH.

GN BCHK.

OS Helicobacillus mobilis.

OC Bacteria; Firmicutes; Clostridia; Clostridiales; Helicobacteriaceae;

OC Helicobacillus.

OX NCBI\_TaxID=28064;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=99061957; PubMed=9843979;

RA Xiong J., Inoue K., Bauer C.E.;

RT "Tracking molecular evolution of photosynthesis by characterization of

RT a major photosynthesis gene cluster from Helicobacillus mobilis.";

RL Proc. Natl. Acad. Sci. U.S.A. 95:14851-14856 (1998).

DR EMBL; AF080002; AAC84033.1; -.

DR InterPro; IPR003672; COBN/Mg chltase.

DR Pfam; PF02514; COBN-Mg chel.; 1.

SQ SEQUENCE 1292 AA; 144854 MW; 323AA0517B07448D CRC64;

Query Match 33.3%; Score 58; DB 2; Length 1292;

Best Local Similarity 42.3%; Pred. No. 16;

Matches 11; Conservative 3; Mismatches 6; Indels 6; Gaps 1;

QY 9 ETRTTPDINPAWYTG-----RGIRPV 28

DB 1176 ETRTKLPNPKWYEGMLKHGVEGVREI 1201

RESULT 7

Q95YJ8

ID Q95YJ8 PRELIMINARY; PRT; 355 AA.

AC Q95YJ8

DT 01-DEC-2001 (TRENBLrel. 19, Created)

DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)

DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)

DE Zic related protein 1a.

GN CS-ZICR1A.

OS Ciona savignyi.

OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;

OC Phlebobranchia; Cionidae; Ciona.

OX NCBI\_TaxID=51511;

RN [1]

RP SEQUENCE FROM N.A.

RA Imai K.S., Satou Y.;

RT "Ciona savignyi genes.";

RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB057747; BAB68356.1; -.

DR InterPro; IPR007087; Znf C2H2.

DR Pfam; PF00096; Zf-C2H2; 4.

DR SMART; SM00355; Znf C2H2; 4.

DR PROSITE; PS00028; ZINC FINGER C2H2\_1; 3.

DR PROSITE; PS0157; ZINC FINGER C2H2\_2; 4.

KW Metal-binding; Zinc; Zinc-finger.

SQ SEQUENCE 355 AA; 40876 MW; E58F5DEDD812E8AC CRC64;

Query Match

Best Local Similarity 32.2%; Score 56; DB 5; Length 355;

Matches 11; Conservative 8; Mismatches 7; Indels 4; Gaps 1;

QY 2 RAHQSHMETRTPDIN----PAWYTGIRPV 27

DB 27 QAHQSHSDSKPQLNSVPSAAYYAGYGWIP 56

RESULT 8

Q9UJF9

ID Q9UJF9 PRELIMINARY; PRT; 54 AA.

AC Q9UJF9

DT 01-MAY-2000 (TRENBLrel. 13, Created)

DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)

DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)

DE DJ479J7.3 (Sushi-repeat protein (SRPUL)) (Fragment).

GN DJ479J7.3.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Lawlor S.;

RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AL035608; CAB55682.1; -.

FT NON TER 54

SQ SEQUENCE 54 AA; 6110 MW; E2F3C39F7B961A9F CRC64;

Query Match

Best Local Similarity 30.5%; Score 53; DB 4; Length 54;

Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 12 TPDINPAWYTGIRPV 27

DB 18 TPAVPTWYAGSGYYP 33

RESULT 9

Q8BHP9

ID Q8BHP9 PRELIMINARY; PRT; 139 AA.

AC Q8BHP9

DT 01-MAR-2003 (TRENBLrel. 23, Created)

DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)

DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)

DE Weakly similar to hypothetical protein KIAA0574.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Body;

RX MEDLINE=22354683; PubMed=12466851;

RA The FANTOM Consortium,

RA The RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs.";

RL Nature 420:563-573 (2002).

DR EMBL; AK017751; BAC25529.1; -.

KW Hypothetical protein.

SQ SEQUENCE 139 AA; 14740 MW; A08DD1B09441B259 CRC64;

Query Match

Best Local Similarity 30.5%; Score 53; DB 11; Length 139;

Matches 12; Conservative 2; Mismatches 6; Indels 4; Gaps 2;

QY 4 HQHSMETRTPDINPAWYTGIRPV 27

DB 47 HQHS-STGDDP---TWKTGQGTKP 66

## RESULT 10

OS Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;  
 OC Prevotellaceae; Prevotella.  
 OX NCBI\_TaxID=77768;  
 RN [1]  
 RP SEQUENCE OF 9-191 FROM N.A.  
 RC STRAIN=W384;  
 RA Walker N.D., McEwan N.R., Wallace R.J.;  
 RT "Prevotella albensis putative uvra gene";  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.  
 DR EMBL; AF070990; AAC24131.2; -;  
 DR InterPro; IPR003439; ABC transporter.  
 DR ProDom; PD000006; ABC transporter; 1.  
 DR PROSITE; PS00211; ABC TRANSPORTER; 1.  
 KM ATP-binding; Transport.  
 FT NON\_TER 1  
 FT NON\_TER 223  
 SQ SEQUENCE 223 AA; 24579 MW; 2563BA0B0C2996C9 CRC64;

Query Match 30.5%; Score 53; DB 2; Length 223;  
 Best Local Similarity 39.5%; Pred. No. 12;  
 Matches 15; Conservative 1; Mismatches 4; Indels 18; Gaps 2;

QY 11 RTPDINPAWYTG-----RGIRPVGRF 31  
 ||| ||| |||  
 DB 56 RTPRNPATYGVFSDIRTLFVGLPEAKIRGYKP-GRF 92

## RESULT 11

OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kuroiwa H., Inukai T., Inaba T., Goi K., Chang K.-S., Siniyo T.,  
 RA Rakestraw K.M., Naeve C.W., Look T.A.;  
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
 CC [2]  
 RP SEQUENCE FROM N.A.  
 RA Huang C.-H., Chen H., Peng J., Chen Y.;  
 RT "Cloning and characterization of the sushi-repeat containing protein  
 (SRP) as a novel interaction partner of Rh type C glycoprotein  
 (RhCG).";  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF060567; AAC15765.1; -;  
 DR EMBL; AF393649; AAM73693.1; -;  
 DR InterPro; IPR001128; Cytochrome\_P450.  
 DR InterPro; IPR003410; Hyalin.  
 DR InterPro; IPR000436; Sushi\_SCR\_CCP.  
 DR Pfam; PF02494; HYR; 1.  
 DR PROSITE; PS00084; sushi; 3.  
 DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
 SQ SEQUENCE 465 AA; 52971 MW; 4D752B187FF3EFB8 CRC64;

Query Match 30.5%; Score 53; DB 2; Length 223;  
 Best Local Similarity 39.5%; Pred. No. 12;  
 Matches 15; Conservative 1; Mismatches 4; Indels 18; Gaps 2;

Query Match 30.5%; Score 53; DB 4; Length 465;  
 Best Local Similarity 50.0%; Pred. No. 27;  
 Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 12 TPDINPAWYTGIRP 27  
 ||| : ||| |||  
 DB 18 TPAVPTWYAGSGYYP 33

## RESULT 12

OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC Straussberg R.;  
 RA Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC020733; AAH20733.1; -;  
 DR InterPro; IPR001128; Cytochrome\_P450.  
 DR InterPro; IPR003410; Hyalin.  
 DR InterPro; IPR000436; Sushi\_SCR\_CCP.  
 DR Pfam; PF02494; HYR; 1.  
 DR Pfam; PF00084; sushi; 3.  
 DR SMART; SM00032; CCP; 3.  
 DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
 SQ SEQUENCE 465 AA; 52957 MW; 3D7229487DA1B8BD CRC64;

Query Match 30.5%; Score 53; DB 4; Length 465;  
 Best Local Similarity 50.0%; Pred. No. 27;  
 Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 12 TPDINPAWYTGIRP 27  
 ||| : ||| |||  
 DB 18 TPAVPTWYAGSGYYP 33

## RESULT 13

OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Straussberg R.;  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.  
 DR EMBL; BC024826; AAH24826.1; -;  
 DR InterPro; IPR002059; Cold\_shock.  
 DR InterPro; IPR002016; Peroxidase.  
 DR Pfam; PF00313; CSD; 7.  
 DR ProDom; PD000621; Cold\_shock; 1.  
 DR SMART; SM00357; CSP; 4.  
 DR PROSITE; PS00352; COLD\_SHOCK; 3.  
 DR PROSITE; PS00436; PEROXIDASE\_2; 1.  
 KW Hypothetical protein.  
 FT NON\_TER 1  
 FT NON\_TER 1

Query Match 30.5%; Score 53; DB 4; Length 465;  
 Best Local Similarity 50.0%; Pred. No. 27;  
 Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;



```

RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.
DR EMBL; BC016898; AAH16898.1; -.
DR MGD; MGI:22356; D3Jfr1.
DR InterPro; IPR002059; Cold shock.
DR InterPro; IPR002016; Peroxidase.
DR Pfam; PF00313; CSD; 8.
DR ProDom; PD000621; Cold shock; 1.
DR SMART; SM00357; CSP; 5.
DR PROSITE; PS00352; COLD SHOCK; 4.
DR PROSITE; PS00436; PEROXIDASE_2; 2.
KW Hypothetical protein.
SQ SEQUENCE 798 AA; 88790 MW; 731065F734C60009 CRC64;

Query Match 30.5%; Score 53; DB 11; Length 798;
Best Local Similarity 43.5%; Pred. No. 50;
Matches 10; Conservative 5; Mismatches 8; Indels 0;

Qy 6 HSMETRTDPINPAWYTGIRPV 28
||: ||: ||: ||: ||: ||:
Db 583 HSVNGITEANPTIYSGKVRPL 605

Search completed: December 3, 2003, 19:14:01
Job time : 26.5 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: December 3, 2003, 18:57:14 ; Search time 33.75 Seconds  
(without alignments)  
145.793 Million cell updates/sec

Title: US-09-868-885b-32  
Perfect score: 171  
Sequence: 1 SRTHSHMEIRTPDINPAWYASGRIPVGRF 31

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_19Jun03:\*  
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*  
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*  
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Match | Length | DB ID       | Description        |
|------------|-------|-------|--------|-------------|--------------------|
| 1          | 171   | 100.0 | 31     | 18 AAW31391 | Human type G prote |
| 2          | 171   | 100.0 | 31     | 20 AAW97235 | Human type ligand  |
| 3          | 171   | 100.0 | 31     | 20 AAW87615 | Human 19P2 ligand  |
| 4          | 171   | 100.0 | 31     | 21 AAB10362 | Human oxytocin sec |
| 5          | 171   | 100.0 | 31     | 21 AAY49291 | 19P2 ligand peptid |
| 6          | 171   | 100.0 | 31     | 22 AAG62531 | Human CRH releasin |
| 7          | 171   | 100.0 | 31     | 22 AAB90991 | Prolactin releasin |
| 8          | 171   | 100.0 | 31     | 23 AAE26401 | Human PrRP-31 pep  |
| 9          | 171   | 100.0 | 31     | 24 ABU60827 | Peptide production |

|    |     |       |    |             |                    |
|----|-----|-------|----|-------------|--------------------|
| 10 | 171 | 100.0 | 31 | 24 ABU60843 | Peptide production |
| 11 | 171 | 100.0 | 32 | 18 AAW31392 | Human type G prote |
| 12 | 171 | 100.0 | 32 | 21 AAB10363 | Human oxytocin sec |
| 13 | 171 | 100.0 | 32 | 22 AAG62532 | Human CRH releasin |
| 14 | 171 | 100.0 | 32 | 24 ABU60844 | Peptide production |
| 15 | 171 | 100.0 | 33 | 18 AAW31393 | Human type G prote |
| 16 | 171 | 100.0 | 33 | 21 AAB10364 | Human oxytocin sec |
| 17 | 171 | 100.0 | 33 | 22 AAG62533 | Human CRH releasin |
| 18 | 171 | 100.0 | 33 | 24 ABU60845 | Peptide production |
| 19 | 171 | 100.0 | 87 | 18 AAW31390 | Human type G prote |
| 20 | 171 | 100.0 | 87 | 20 AAW97226 | Human type ligand  |
| 21 | 171 | 100.0 | 87 | 21 AAB10361 | Human oxytocin sec |
| 22 | 171 | 100.0 | 87 | 22 AAG62530 | Human CRH releasin |
| 23 | 165 | 96.5  | 30 | 21 AAY49299 | 19P2 ligand peptid |
| 24 | 162 | 94.7  | 31 | 22 AAB90995 | Prolactin releasin |
| 25 | 158 | 92.4  | 31 | 18 AAW31371 | Bovine G protein-c |
| 26 | 158 | 92.4  | 31 | 20 AAW97218 | Bovine pituitary-d |
| 27 | 158 | 92.4  | 31 | 20 AAW87613 | Bovine 19P2 ligand |
| 28 | 158 | 92.4  | 31 | 20 AAW95188 | Bovine pituitary-d |
| 29 | 158 | 92.4  | 31 | 21 AAB10347 | Bovine oxytocin se |
| 30 | 158 | 92.4  | 31 | 21 AAY49290 | 19P2 ligand peptid |
| 31 | 158 | 92.4  | 31 | 21 AAY49298 | 19P2 ligand peptid |
| 32 | 158 | 92.4  | 31 | 22 AAG62516 | Bovine CRH releasi |
| 33 | 158 | 92.4  | 31 | 23 AAE26399 | Bovine PrRP-31 pep |
| 34 | 158 | 92.4  | 31 | 24 ABU60825 | Peptide production |
| 35 | 158 | 92.4  | 31 | 24 ABU60831 | Peptide production |
| 36 | 158 | 92.4  | 32 | 18 AAW31372 | Bovine G protein-c |
| 37 | 158 | 92.4  | 32 | 20 AAW95189 | Bovine pituitary-d |
| 38 | 158 | 92.4  | 32 | 21 AAB10348 | Bovine oxytocin se |
| 39 | 158 | 92.4  | 32 | 22 AAG62517 | Bovine CRH releasi |
| 40 | 158 | 92.4  | 32 | 24 ABU60832 | Peptide production |
| 41 | 158 | 92.4  | 33 | 18 AAW31373 | Bovine G protein-c |
| 42 | 158 | 92.4  | 33 | 20 AAW95190 | Bovine pituitary-d |
| 43 | 158 | 92.4  | 33 | 21 AAB10349 | Bovine oxytocin se |
| 44 | 158 | 92.4  | 33 | 21 AAY49297 | 19P2 ligand peptid |
| 45 | 158 | 92.4  | 33 | 22 AAG62518 | Bovine CRH releasi |

## ALIGNMENTS

### RESULT 1

AAW31391  
ID AAW31391 standard; Peptide; 31 AA.

AC AAW31391;

DT 06-APR-1998 (first entry)

DE Human type G protein-coupled receptor ligand fragment 1.

DE G protein-coupled receptor; ligand binding; pharmaceutical;  
modulator; pituitary; central nervous system; pancreas; prophylactic;  
therapeutic agent.

OS Homo sapiens.

XX WO9724436-A2.

XX 10-JUL-1997.

XX 26-DEC-1996; 96WO-JP03821.

XX 18-SEP-1996; 96JP-0246573.

XX 28-DEC-1995; 95JP-0343371.

XX 15-MAR-1996; 96JP-0059419.

XX 12-AUG-1996; 96JP-0211805.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;  
Kawamata Y, Kitada C;

DR WPI; 1997-363672/33.  
 DR N-PSDB; AAV02428.  
 PT Ligand peptide for G protein-coupled receptor - acts by modulating  
 PT function in the central nervous system, pancreas and pituitary gland  
 XX  
 XX Claim 2; Page 184; 258pp; English.  
 CC This sequence represents a peptide fragment from a novel human type  
 CC ligand polypeptide corresponding to amino acid residues 23 to 53 of the  
 CC sequence represented in AAW1390 and is used in an assay to monitor  
 CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
 CC compositions containing this ligand may be used as a pituitary function  
 CC modulator, a central nervous system modulator or a pancreatic function  
 CC modulator. This ligand could have specific applications as a  
 CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
 CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
 CC trauma, growth hormone secretory disease, hyper- and polyphagia,  
 CC hyperlipidaemia, hypercholesterolaemia, hyperglycaemia,  
 CC hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease,  
 CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
 CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
 CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
 CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
 CC oligogalactia. Assays can also be developed to screen compounds which are  
 CC capable of altering the binding activity of the ligand affecting  
 CC activation of the G protein-coupled receptor protein.  
 XX  
 XX Sequence 31 AA;  
 Query Match 100.0%; Score 171; DB 18; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 5.8e-19;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SRTRHSMEIRTPDINPAWYASRGIRPVGRF 31  
 DB 1 SRTRHSMEIRTPDINPAWYASRGIRPVGRF 31  
 RESULT 2  
 AAW97235  
 ID AAW97235 standard; peptide; 31 AA.  
 AC AAW97235;  
 XX  
 XX 06-MAY-1999 (first entry)  
 DT  
 XX  
 XX Human type ligand polypeptide fragment.  
 DE  
 XX Rat type ligand; modulation; prolactin secretion;  
 KW G protein-coupled receptor; GPCR; hypocoovarianism; gonocyst cacogenesis;  
 KW menopausal syndrome; euthyroid; hypometabolism; lactation;  
 KW pituitary adenomatosis; brain tumour; emmenopathy; autoimmune disease;  
 KW prolactinoma; infertility; impotence; amenorrhea; galactorrhea;  
 KW acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;  
 KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;  
 KW contraceptive; placental function; choriocarcinoma; hydatid mole;  
 KW irruption mole; abortion; unthrifty fetus; abnormal saccharometabolism;  
 KW abnormal lipidmetabolism; oxytocia.  
 XX  
 XX Homo sapiens.  
 OS  
 XX WO9858962-A1.  
 PN  
 XX 30-DEC-1998.  
 PD  
 XX 22-JUN-1998; 98WO-JP02765.  
 XX  
 XX 23-JUN-1997; 97JP-0165437.  
 XX  
 XX (TAKE ) TAKEDA CHEM IND LTD.  
 PA  
 XX Fujii R, Hinuma S, Kawamata Y, Matsumoto H;  
 PI

XX WPI; 1999-105614/09.  
 XX Use of G protein-coupled receptor ligands - for modulating prolactin  
 PT secretion or placental function, e.g. for treating menopausal  
 PT syndrome, tumours, autoimmune disease or abnormal pregnancy  
 XX  
 XX Claim 3; Page 159; 241pp; English.  
 CC The present sequence represents a human type ligand fragment. It  
 CC is used in the course of the invention. The specification describes  
 CC an agent for modulating prolactin secretion which comprises a  
 CC ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)  
 CC protein. The agents for promoting prolactin secretion can be used for  
 CC treating or preventing hypocoovarianism, gonocyst cacogenesis, menopausal  
 CC syndrome, euthyroid or hypometabolism. They can be used for promoting  
 CC lactation in a domestic mammal and as an aphrodisiac. The agents for  
 CC inhibiting prolactin secretion can be used for treating or preventing  
 CC pituitary adenomatosis, brain tumour, emmenopathy, autoimmune disease,  
 CC prolactinoma, infertility, impotence, amenorrhea, galactorrhea,  
 CC acromegaly, Chiari-Frommel syndrome, Argonz-del Castillo syndrome,  
 CC Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dyszoospermia.  
 CC The inhibitory agents can also be used as contraceptives. The agents for  
 CC modulating placental function can be used for treating or preventing  
 CC choriocarcinoma, hydatid mole, irruption mole, abortion, unthrifty fetus,  
 CC abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.  
 XX  
 XX Sequence 31 AA;  
 Query Match 100.0%; Score 171; DB 20; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 5.8e-19;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SRTRHSMEIRTPDINPAWYASRGIRPVGRF 31  
 DB 1 SRTRHSMEIRTPDINPAWYASRGIRPVGRF 31  
 RESULT 3  
 AAW87615  
 ID AAW87615 standard; Peptide; 31 AA.  
 AC AAW87615;  
 XX  
 XX 29-MAR-1999 (first entry)  
 DT  
 XX  
 XX Human 19p2 ligand.  
 DE  
 XX 19p2 ligand; G protein coupled receptor; pituitary;  
 KW prolactin releasing peptide; human; dementia; breast cancer;  
 KW therapy.  
 XX  
 XX Homo sapiens.  
 OS  
 XX EP887417-A2.  
 PN  
 XX 30-DEC-1998.  
 PD  
 XX 25-JUN-1998; 98EP-0111725.  
 XX  
 XX 27-JUN-1997; 97JP-0172118.  
 XX  
 XX (TAKE ) TAKEDA CHEM IND LTD.  
 PA  
 XX Moriya T, Nishimura O, Suenaga M, Tanaka Y;  
 PI  
 XX WPI; 1999-047884/05.  
 DR  
 XX Producing a 19p2 pituitary G protein receptor ligand - by cleavage  
 PT of a fusion protein, useful for preventing and treating dementia,  
 PT breast cancer, renal failure and autoimmune disease  
 XX  
 XX Claim 5; Page 35; 56pp; English.  
 PS

XX This is the amino acid sequence of the human pituitary G  
 CC protein-coupled receptor ligand 19P2L. A method suitable for  
 CC commercial high-level production of 19P2L comprises expressing  
 CC the ligand in host cells as a recombinant fusion protein e.g. with  
 CC human basic fibroblast growth factor (see AAV83796-97) that has  
 CC been modified to include an N-terminal cysteine residue. The  
 CC ligand is released from the fusion by cyanylation followed by  
 CC ammonolysis. 19P2L has prolactin secretion-stimulating and (at  
 CC high doses) prolactin secretion-inhibiting properties. It can be  
 CC used in the treatment and prevention of various diseases including:  
 CC senile dementia, cerebrovascular dementia, and dementia associated  
 CC with: neurological disorders (e.g. Alzheimer's disease, Parkinson's  
 CC disease, Pick's disease, Huntington's disease), infectious diseases  
 CC (e.g. Creutzfeldt-Jakob's), endocrine or metabolic disease or  
 CC toxicosis (e.g. hypothyroidism, vitamin B12 deficiency, alcoholism,  
 CC intoxication by drugs, metal and organic compounds), tumourigenic  
 CC diseases (e.g. brain tumour), traumatic diseases (e.g. chronic  
 CC subarachnoid haemorrhage), and other types of dementia, depression,  
 CC hyperactive child syndrome (microencephalopathy) and disturbance of  
 CC consciousness. It is also useful for prevention and treatment of  
 CC diseases associated with prolactin hypo and hypersecretion  
 CC respectively, including: hyperprolactinaemia, pituitary adenoma,  
 CC breast cancer, infertility, impotence and autoimmune disease  
 CC (hypersecretion disorders), and seminal vesicle hypoplasia,  
 CC osteoporosis, menopausal syndrome and renal failure (hyposecretion  
 CC disorders). The 19P2 polypeptide/amide is also useful as a test  
 CC reagent for study of the prolactin secretory function or as a  
 CC lactogogue in mammalian farm animals.

XX SQ Sequence 31 AA;  
 Query Match 100.0%; Score 171; DB 20; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 5.8e-19;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRTRHSHMEIRTPDINPAWYASRGIRPVGRF 31  
 DB 1 SRTRHSHMEIRTPDINPAWYASRGIRPVGRF 31

RESULT 4  
 ID AAB10362 standard; peptide; 31 AA.  
 AC AAB10362;  
 XX 24-NOV-2000 (first entry)  
 DT Human oxytocin secretion promoting peptide SEQ ID NO: 32.  
 DE Human; oxytocin secretion promoter; G protein-coupled receptor protein;  
 KW treatment; disease; pain; atonic bleeding; uterine recovery failure;  
 KW caesarean section; artificial fertilization; galactostasis; goat; pig;  
 KW veterinary medicine; milk production.  
 XX Homo sapiens.  
 OS WO200038704-A1.  
 FN 06-JUL-2000.  
 PD 22-DEC-1999; 99WO-JP07199.  
 XX 25-DEC-1998; 98JP-0369585.  
 PR (TAKE ) TAKEDA CHEM IND LTD.  
 PA Matsumoto H, Kitada C, Hinuma S;  
 XX WPI; 2000-452298/39.  
 DR Physiologically-active polypeptide recognized as ligand by G

PT protein-coupled receptor protein, for promoting secretion of oxytocin,  
 PT as drugs for diseases relating to oxytocin secretion and in veterinary  
 PT medicine -  
 XX Disclosure; Page 62; 72pp; Japanese.  
 XX This invention describes a novel oxytocin secretion-regulating agent  
 CC which contains a ligand peptide or its salt for the G protein-coupled  
 CC receptor protein. It is useful in the form of drugs for ameliorating,  
 CC preventing and treating diseases relating to oxytocin secretion e.g.  
 CC weak pains and atonic bleeding, before and after expulsion of placenta,  
 CC uterine recovery failure, caesarean section, stoppage of artificial  
 CC fertilization or galactostasis and is also applicable in veterinary  
 CC medicine for promoting milk production in cow, goat and pig. This  
 CC sequence represents a human peptide which acts as an oxytocin secretion  
 CC promoter.

XX SQ Sequence 31 AA;  
 Query Match 100.0%; Score 171; DB 21; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 5.8e-19;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRTRHSHMEIRTPDINPAWYASRGIRPVGRF 31  
 DB 1 SRTRHSHMEIRTPDINPAWYASRGIRPVGRF 31

RESULT 5  
 ID AAY49291 standard; peptide; 31 AA.  
 XX AAY49291;  
 AC AAY49291;  
 XX 22-FEB-2000 (first entry)  
 DT 19P2 ligand peptide fragment.  
 DE Monoclonal antibody; 19P2 ligand; diagnosis; prolactin secretion;  
 KW pituitary; regulatory mechanism; central nervous system; pancreatic.  
 XX Homo sapiens.  
 OS Key Location/Qualifiers  
 FH Modified-site 31 /note= "C-terminal amide"  
 FT WO9960112-A1.  
 XX 25-NOV-1999.  
 PD 20-MAY-1999; 99WO-JP02650.  
 XX 21-MAY-1998; 98JP-0140293.  
 PR (TAKE ) TAKEDA CHEM IND LTD.  
 PA Matsumoto H, Kitada C, Hinuma S;  
 XX WPI; 2000-039381/03.  
 DR New monoclonal antibodies, useful in diagnosis, as drugs and in  
 PT studying diseases related to ligand abnormality -  
 XX Disclosure; Page 26; 73pp; Japanese.  
 XX The invention provides a monoclonal antibody which has a specific  
 CC reaction with the part peptide of the C-terminal of 19P2 ligand or its  
 CC derivative. The antibodies can be used in diagnosis or to treat or  
 CC prevent diseases associated with abnormality in the pituitary function  
 CC regulatory mechanism (e.g. promotion of prolactin secretion), central  
 CC nervous regulatory mechanism, and pancreatic function regulatory  
 CC mechanism. The antibody-based immunoassay can also be applied in

CC clarifying the physiological functions of the ligand and its derivative.  
 CC Sequences AAY4290-302 represent peptide fragments of the 19P2 ligand.

XX  
 SQ Sequence 31 AA;  
 Query Match 100.0%; Score 171; DB 21; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 5.8e-19;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRTHSHMEIRTPDINPAWYASRGIRPVGRF 31  
 |||||  
 Db 1 SRTHSHMEIRTPDINPAWYASRGIRPVGRF 31

## RESULT 6

AA62531  
 ID AAG62531 standard; peptide; 31 AA.

XX  
 AC AAG62531;

XX  
 DT 24-AUG-2001 (first entry)

XX Human CRH releasing protein related peptide SEQ ID NO: 32.

XX Human; corticotrophin releasing hormone; CRH; G protein receptor ligand;  
 KW analgesic; hyperaldosteronism; hypercortisolemia; hypoadrenocorticism;  
 KW Addison's disease; adrenal gland hyperfunction; obesity.

XX Homo sapiens.

XX WO200135984-A1.

XX  
 PD 25-MAY-2001.

XX  
 PF 17-NOV-2000; 2000WO-JP08119.

XX  
 PR 18-NOV-1999; 99JP-0327900.

XX  
 PR 26-SEP-2000; 2000JP-0297073.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Kitada C, Matsumoto H, Hinuma S;

XX WPI; 2001-355552/37.

XX Use of G protein receptor ligand or peptide for controlling  
 PT corticotropin releasing hormone secretion -

XX Claim 3; Page 73-74; 90pp; Japanese.

XX The present sequence describes a method of controlling the secretion of  
 CC corticotrophin releasing hormone (CRH), involving the use of a G protein  
 CC receptor ligand. This can be used to control the secretion of CRH and is  
 CC useful as an analgesic or for treating, preventing or ameliorating  
 CC diseases associated with CRH secretion such as hyperaldosteronism,  
 CC hypercortisolemia, secondary or chronic hypoadrenocorticism, Addison's  
 CC disease (including, boredom, nausea, pigmentation, hypogonadism, hair  
 CC loss, and hypotension), adrenal gland hypofunction and obesity. The  
 CC present sequence is a peptide used in the exemplification of the  
 CC invention.

XX Sequence 31 AA;

Query Match 100.0%; Score 171; DB 22; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 5.8e-19;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRTHSHMEIRTPDINPAWYASRGIRPVGRF 31  
 |||||

Db 1 SRTHSHMEIRTPDINPAWYASRGIRPVGRF 31

## RESULT 7

AA690991

ID AAB90991 standard; peptide; 31 AA.

XX  
 AC AAB90991;

XX  
 DT 22-JUN-2001 (first entry)

XX Prolactin releasing peptide SEQ ID NO:165.

XX Protection; endogenous therapeutic peptide; peptidase; conjugation;  
 KW blood component; modification; succinimide; maleimide group; amino;  
 KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.

XX Homo sapiens.

XX Synthetic.

XX WO200069900-A2.

XX  
 PD 23-NOV-2000.

XX  
 PF 17-MAY-2000; 2000WO-US13576.

XX  
 PR 17-MAY-1999; 99US-0134406.

XX  
 PR 10-SEP-1999; 99US-0153406.

XX  
 PR 15-OCT-1999; 99US-0159783.

XX (CONJ-) CONJUCHEM INC.

XX Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;

XX WPI; 2001-112059/12.

XX Modifying and attaching therapeutic peptides to albumin prevents  
 PT peptidase degradation, useful for increasing length of in vivo activity

XX Disclosure; Page 244; 733pp; English.

XX The present invention describes a modified therapeutic peptide (I)  
 CC comprising a therapeutically active amino acid region (III) and a  
 CC reactive group (II) (e.g. succinimide and maleimide groups) attached to  
 CC a less therapeutically active amino acid region (IV), which covalently  
 CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
 CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
 CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
 CC factors and neurotransmitters, to protect them from peptidase activity  
 CC in vivo for the treatment of various disorders. Endogenous therapeutic  
 CC peptides are not suitable as drug candidates as they require frequent  
 CC administration due to rapid degradation by peptidases in the body.  
 CC Modifying and attaching therapeutic peptides to albumin prevents or  
 CC reduces the action of peptidases to increase length of activity (half  
 CC life) and specificity as bonding to large molecules decreases  
 CC intracellular uptake and interference with physiological processes.  
 CC AAB90829 to AAB92441 represent peptides which can be used in the  
 CC exemplification of the present invention.

XX Sequence 31 AA;

Query Match 100.0%; Score 171; DB 22; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 5.8e-19;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRTHSHMEIRTPDINPAWYASRGIRPVGRF 31  
 |||||

Db 1 SRTHSHMEIRTPDINPAWYASRGIRPVGRF 31

## RESULT 8

AAE26401

ID AAE26401 standard; peptide; 31 AA.

XX  
 AC AAE26401;

XX

DT 13-DEC-2002 (first entry)  
 XX Human PrRP-31 peptide.  
 DE  
 XX  
 KW Human; wakefulness; sleep disorder; prolactin releasing peptide receptor;  
 KW PrRP; GPR10; therapy; epilepsy; narcolepsy; sleepiness; sleep apnoea;  
 KW insomnia; idiopathic hypersomnia; psychogenic hypersomnia; seizure;  
 KW anticonvulsant.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2002037533-A1.  
 XX  
 PD 28-MAR-2002.  
 XX  
 PF 17-AUG-2001; 2001US-0932161.  
 XX  
 PR 28-APR-2000; 2000US-0560915.  
 XX  
 PA (CIVE/) CIVELLI O.  
 PA (LINS/) LIN S.  
 XX  
 PI Civelli O, Lin S;  
 XX  
 DR WPI; 2002-403931/43.  
 XX  
 PT Screening for compounds useful for promoting wakefulness or sleep, and  
 PT for treating sleeping disorders, e.g. insomnia, hypersomnia or sleep  
 PT apnea, comprises administering a prolactin releasing peptide agonist or  
 PT antagonist -  
 XX  
 PS Disclosure; Page 24; 35pp; English.  
 XX  
 CC The present invention relates to a method of screening for compounds  
 CC for promoting wakefulness or sleep in a mammal. The method involves  
 CC administering a prolactin releasing peptide (PrRP) receptor (GPR10)  
 CC agonist or antagonist respectively and determining the ability of the  
 CC compound to promote wakefulness or sleep. The compounds identified  
 CC from the method are used in the therapy of epilepsy and other diseases  
 CC associated with absence seizures and in promoting wakefulness and sleep  
 CC in individuals having sleep disorders such as insomnia and narcolepsy.  
 CC PrRP receptor agonists may be used to treat common disorders which lead  
 CC to sleepiness, e.g. sleep apnoea, narcolepsy, idiopathic hypersomnia  
 CC and psychogenic hypersomnia. PrRP receptor antagonists are useful for  
 CC promoting sleep and for treating insomnia such as adjustment sleep  
 CC disorder and psychophysiological insomnia. The present sequence is  
 CC human PrRP-31 peptide.  
 XX  
 SQ Sequence 31 AA;  
 Query Match 100.0%; Score 171; DB 23; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 5.8e-19;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SRTRHSHMEIRTPDINPAWYASRGIRPVGRF 31  
 |||||  
 DB 1 SRTRHSHMEIRTPDINPAWYASRGIRPVGRF 31  
 |||||  
 RESULT 9  
 ABU60827  
 ID ABU60827 standard; Peptide; 31 AA.  
 XX  
 AC ABU60827;  
 XX  
 DT 06-MAY-2003 (first entry)  
 XX  
 DE Peptide production by gene recombination associated peptide #11.  
 KW Peptide production; low-molecular peptide; KiSS-1; GPR8 ligand;  
 KW gene recombination.  
 XX  
 OS Homo sapiens.

XX WO200292829-A1.  
 PN  
 XX  
 PD 21-NOV-2002.  
 XX  
 PP 16-MAY-2002; 2002WO-JP04735.  
 XX  
 PR 17-MAY-2001; 2001JP-0147341.  
 XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PI Nishimura O, Suenaga M, Ito T, Kitada C;  
 XX  
 DR WPI; 2003-129302/12.  
 XX  
 PT Process for producing peptides e.g. KiSS-1 peptide and GPR8 ligand for  
 PT subsequent applications by gene recombination technique through tandem  
 PT repeats to provide precursor protein with specific cleavage sites -  
 XX  
 PS Disclosure; Page 59; 87pp; Japanese.  
 XX  
 CC The invention describes a method of producing a peptide comprising the  
 CC excision of the N and C-terminals of a target peptide with enzymes or  
 CC chemically through the attached cleavage sites repeated by ligation in a  
 CC precursor protein. The method is for producing (low-molecular) peptides  
 CC e.g. KiSS-1 peptide and GPR8 ligand for subsequent applications by the  
 CC gene recombination technique through tandem repeats to provide  
 CC a precursor protein with specific cleavage sites. With this method,  
 CC peptide production can be carried out easily to provide large quantities  
 CC of the required peptides. This is the amino acid sequence of a peptide  
 CC associated with the peptide production method of the invention.  
 XX  
 SQ Sequence 31 AA;  
 Query Match 100.0%; Score 171; DB 24; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 5.8e-19;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SRTRHSHMEIRTPDINPAWYASRGIRPVGRF 31  
 |||||  
 DB 1 SRTRHSHMEIRTPDINPAWYASRGIRPVGRF 31  
 |||||  
 RESULT 10  
 ABU60843  
 ID ABU60843 standard; Peptide; 31 AA.  
 XX  
 AC ABU60843;  
 XX  
 DT 06-MAY-2003 (first entry)  
 XX  
 DE Peptide production by gene recombination associated peptide #27.  
 KW Peptide production; low-molecular peptide; KiSS-1; GPR8 ligand;  
 KW gene recombination.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200292829-A1.  
 XX  
 PD 21-NOV-2002.  
 XX  
 PP 16-MAY-2002; 2002WO-JP04735.  
 XX  
 PR 17-MAY-2001; 2001JP-0147341.  
 XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PI Nishimura O, Suenaga M, Ito T, Kitada C;  
 XX  
 DR WPI; 2003-129302/12.  
 XX  
 PT Process for producing peptides e.g. KiSS-1 peptide and GPR8 ligand for

PT subsequent applications by gene recombination technique through tandem  
 PT repeats to provide precursor protein with specific cleavage sites -  
 XX  
 PS Disclosure; Page 67; 87pp; Japanese.  
 XX  
 CC The invention describes a method of producing a peptide comprising the  
 CC excision of the N and C-terminals of a target peptide with enzymes or  
 CC chemically through the attached cleavage sites repeated by ligation in a  
 CC precursor protein. The method is for producing (low-molecular) peptides  
 CC e.g. KiSS-1 peptide and GPR8 ligand for subsequent applications by the  
 CC gene recombination technique through tandem repeats to provide  
 CC a precursor protein with specific cleavage sites. With this method,  
 CC peptide production can be carried out easily to provide large quantities  
 CC of the required peptides. This is the amino acid sequence of a peptide  
 CC associated with the peptide production method of the invention.  
 XX  
 SQ Sequence 31 AA;  
 Query Match 100.0%; Score 171; DB 24; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 5.8e-19;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SRTHRSMEIRTPDINPAWYASRGIRPVGRF 31  
 DB 1 SRTHRSMEIRTPDINPAWYASRGIRPVGRF 31  
 RESULT 11  
 AAW31392  
 ID AAW31392 standard; Peptide; 32 AA.  
 XX  
 AC AAW31392;  
 XX  
 DT 06-APR-1998 (first entry)  
 XX  
 DE Human type G protein-coupled receptor ligand fragment 2.  
 XX  
 KW G protein-coupled receptor; ligand binding; pharmaceutical;  
 KW modulator; pituitary; central nervous system; pancreas; prophylactic;  
 KW therapeutic agent.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9724436-A2.  
 XX  
 PD 10-JUL-1997.  
 XX  
 PF 26-DEC-1996; 96WO-JP03821.  
 XX  
 PR 18-SEP-1996; 96JP-0246573.  
 PR 28-DEC-1995; 95JP-0343371.  
 PR 15-MAR-1996; 96JP-0059419.  
 PR 12-AUG-1996; 96JP-0211805.  
 XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;  
 PI Kawanata Y, Kitada C;  
 XX  
 DR WPI; 1997-363672/33.  
 DR N-PSDB; AAV02429.  
 XX  
 PT Ligand peptide for G protein-coupled receptor - acts by modulating  
 PT function in the central nervous system, pancreas and pituitary gland  
 XX  
 PS Claim 2; Page 185; 258pp; English.  
 XX  
 CC This sequence represents a peptide fragment from a novel human type  
 CC ligand polypeptide corresponding to amino acid residues 23 to 54 of the  
 CC sequence represented in AAW31390 and is used in an assay to monitor  
 CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
 CC compositions containing this ligand may be used as a pituitary function  
 CC modulator, a central nervous system modulator or a pancreatic function

CC modulator. This ligand could have specific applications as a  
 CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
 CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
 CC trauma, growth hormone secretory disease, hyper- and polyphagia,  
 CC hyperlipidaemia, hypercholesterolaemia, hyperglycaemia, disease,  
 CC hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease,  
 CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
 CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
 CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
 CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
 CC oligosaccharia. Assays can also be developed to screen compounds which are  
 CC capable of altering the binding activity of the ligand affecting  
 CC activation of the G protein-coupled receptor protein.  
 XX  
 SQ Sequence 32 AA;  
 Query Match 100.0%; Score 171; DB 18; Length 32;  
 Best Local Similarity 100.0%; Pred. No. 6e-19;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SRTHRSMEIRTPDINPAWYASRGIRPVGRF 31  
 DB 1 SRTHRSMEIRTPDINPAWYASRGIRPVGRF 31  
 RESULT 12  
 AAB10363  
 ID AAB10363 standard; peptide; 32 AA.  
 XX  
 AC AAB10363;  
 XX  
 DT 24-NOV-2000 (first entry)  
 XX  
 DE Human oxytocin secretion promoting peptide SEQ ID NO: 33.  
 XX  
 KW Human; oxytocin secretion promoter; G protein-coupled receptor protein;  
 KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
 KW caesarean section; artificial fertilization; galactostasis; goat; pig;  
 KW veterinary medicine; milk production.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200038704-A1.  
 XX  
 PD 06-JUL-2000.  
 XX  
 PF 22-DEC-1999; 99WO-JP07199.  
 XX  
 PR 25-DEC-1998; 98JP-0369585.  
 XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PI Matsumoto H, Kitada C, Hinuma S;  
 XX  
 DR WPI; 2000-452298/39.  
 XX  
 PT Physiologically-active polypeptide recognized as ligand by G  
 PT protein-coupled receptor protein, for promoting secretion of oxytocin,  
 PT as drugs for diseases relating to oxytocin secretion and in veterinary  
 PT medicine -  
 XX  
 PS Disclosure; Page 62; 72pp; Japanese.  
 XX  
 CC This invention describes a novel oxytocin secretion-regulating agent  
 CC which contains a ligand peptide or its salt for the G protein-coupled  
 CC receptor protein. It is useful in the form of drugs for ameliorating,  
 CC preventing and treating diseases relating to oxytocin secretion e.g.  
 CC weak pains and atonic bleeding, before and after expulsion of placenta,  
 CC uterine recovery failure, caesarean section, stoppage of artificial  
 CC fertilization or galactostasis and is also applicable in veterinary  
 CC medicine for promoting milk production in cow, goat and pig. This  
 CC sequence represents a human peptide which acts as an oxytocin secretion  
 CC promoter.

XX SQ Sequence 32 AA;  
Query Match 100.0%; Score 171; DB 21; Length 32;  
Best Local Similarity 100.0%; Pred. No. 6e-19; Indels 0; Gaps 0;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SRTHRSMEIRTPDINPAWYASRGIRPVGRF 31  
DB 1 SRTHRSMEIRTPDINPAWYASRGIRPVGRF 31  
RESULT 13  
AAG62532  
ID AAG62532 standard; peptide; 32 AA.  
AC AAG62532;  
XX  
DT 24-AUG-2001 (first entry)  
XX  
DE Human CRH releasing protein related peptide SEQ ID NO: 33.  
XX  
KW Human; corticotrophin releasing hormone; CRH; G protein receptor ligand;  
KW analgesic; hyperaldosteronism; hypercortisolaemia; hypoadrenocorticism;  
KW Addison's disease; adrenal gland hyperfunction; obesity.  
XX  
OS Homo sapiens.  
XX  
FN WO200135984-A1.  
XX  
PD 25-MAY-2001.  
XX  
PF 17-NOV-2000; 2000WO-JP08119.  
XX  
PR 18-NOV-1999; 99JP-0327900.  
XX  
PR 26-SEP-2000; 2000JP-0297073.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
XX Kitada C, Mateumoto H, Hinuma S;  
XX WPI; 2001-355552/37.  
XX  
PT Use of G protein receptor ligand or peptide for controlling  
PT corticotrophin releasing hormone secretion -  
XX  
PS Disclosure; Page 74; 90pp; Japanese.  
XX  
CC The present sequence describes a method of controlling the secretion of  
CC corticotrophin releasing hormone (CRH), involving the use of a G protein  
CC receptor ligand. This can be used to control the secretion of CRH and is  
CC useful as an analgesic or for treating, preventing or ameliorating  
CC diseases associated with CRH secretion such as hyperaldosteronism,  
CC hypercortisolaemia, secondary or chronic hypoadrenocorticism, Addison's  
CC disease (including boredom, nausea, pigmentation, hypogonadism, hair  
CC loss, and hypotension), adrenal gland hypofunction and obesity. The  
CC present sequence is a peptide used in the exemplification of the  
XX invention.  
XX SQ Sequence 32 AA;  
Query Match 100.0%; Score 171; DB 22; Length 32;  
Best Local Similarity 100.0%; Pred. No. 6e-19; Indels 0; Gaps 0;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SRTHRSMEIRTPDINPAWYASRGIRPVGRF 31  
DB 1 SRTHRSMEIRTPDINPAWYASRGIRPVGRF 31  
RESULT 14  
ABU60844  
ID ABU60844 standard; Peptide; 32 AA.

XX AC ABU60844;  
XX  
DT 06-MAY-2003 (first entry)  
XX  
DE Peptide production by gene recombination associated peptide #28.  
XX  
KW Peptide production; low-molecular peptide; KiSS-1; GPR8 ligand;  
KW gene recombination.  
XX  
OS Homo sapiens.  
XX  
FN WO200292829-A1.  
XX  
PD 21-NOV-2002.  
XX  
PF 16-MAY-2002; 2002WO-JP04735.  
XX  
PR 17-MAY-2001; 2001JP-0147341.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Nishimura O, Suenaga M, Ito T, Kitada C;  
XX WPI; 2003-129302/12.  
XX  
PT Process for producing peptides e.g. KiSS-1 peptide and GPR8 ligand for  
PT subsequent applications by gene recombination technique through tandem  
PT repeats to provide precursor protein with specific cleavage sites -  
XX  
PS Disclosure; Page 67; 87pp; Japanese.  
XX  
CC The invention describes a method of producing a peptide comprising the  
CC excision of the N and C-terminals of a target peptide with enzymes or  
CC chemically through the attached cleavage sites repeated by ligation in a  
CC precursor protein. The method is for producing (low-molecular) peptides  
CC e.g. KiSS-1 peptide and GPR8 ligand for subsequent applications by the  
CC gene recombination technique through tandem repeats to provide  
CC a precursor protein with specific cleavage sites. With this method,  
CC peptide production can be carried out easily to provide large quantities  
CC of the required peptides. This is the amino acid sequence of a peptide  
CC associated with the peptide production method of the invention.  
XX  
SQ Sequence 32 AA;  
Query Match 100.0%; Score 171; DB 24; Length 32;  
Best Local Similarity 100.0%; Pred. No. 6e-19;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SRTHRSMEIRTPDINPAWYASRGIRPVGRF 31  
DB 1 SRTHRSMEIRTPDINPAWYASRGIRPVGRF 31  
RESULT 15  
AAW31393  
ID AAW31393 standard; Peptide; 33 AA.  
XX  
AC AAW31393;  
XX  
DT 06-APR-1998 (first entry)  
XX  
DE Human type G protein-coupled receptor ligand fragment 3.  
XX  
KW G protein-coupled receptor; ligand binding; pharmaceutical;  
KW modulator; pituitary; central nervous system; pancreas; prophylactic;  
KW therapeutic agent.  
XX  
OS Homo sapiens.  
XX  
FN WO9724436-A2.  
XX  
PD 10-JUL-1997.



```

XX PF 26-DEC-1996; 96WO-JP03821.
XX PR 18-SEP-1996; 96JP-0246573.
XX PR 28-DEC-1995; 95JP-0343371.
XX PR 15-MAR-1996; 96JP-0059419.
XX PR 12-AUG-1996; 96JP-0211805.
XX PA (TAKE ) TAKEDA CHEM IND LTD.
XX PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;
XX PI Kawanata Y, Kitada C;
XX DR WPI; 1997-363672/33.
XX DR N-PSDB; AAV02430.
XX PT Ligand peptide for G protein-coupled receptor - acts by modulating
XX PT function in the central nervous system, pancreas and pituitary gland
XX PS Claim 2; Page 185; 258pp; English.
XX CC This sequence represents a peptide fragment from a novel human type
XX CC ligand polypeptide corresponding to amino acid residues 23 to 55 of the
XX CC sequence represented in AAW31390 and is used in an assay to monitor
XX CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical
XX CC compositions containing this ligand may be used as a pituitary function
XX CC modulator, a central nervous system modulator or a pancreatic function
XX CC modulator. This ligand could have specific applications as a
XX CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic
XX CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,
XX CC trauma, growth hormone secretory disease, hyper- and polyphagia,
XX CC hyperlipidaemia, hypercholesterolaemia, hyperglycaemia,
XX CC hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease,
XX CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,
XX CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,
XX CC acute myocardial infarction, infertility, spinocerebellar degeneration,
XX CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or
XX CC oligogalactia. Assays can also be developed to screen compounds which are
XX CC capable of altering the binding activity of the ligand affecting
XX CC activation of the G protein-coupled receptor protein.
XX SQ Sequence 33 AA;
    Query Match 100.0%; Score 171; DB 18; Length 33;
    Best Local Similarity 100.0%; Pred. No. 6.3e-19;
    Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
    Qy 1 SRTRHSMETPTDINPAWYASRGIRPVGRF 31
    Db 1 SRTRHSMETPTDINPAWYASRGIRPVGRF 31

```

Search completed: December 3, 2003, 19:11:24  
 Job time : 33.75 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 19:09:05 ; Search time 11.5 Seconds  
(without alignments)  
114.055 Million cell updates/sec

Title: US-09-868-885b-32

Perfect score: 171

Sequence: 1 SRTSHSMEIRTPDINPAWYASRGIRPVGRF 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/prodata/1/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/prodata/1/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/prodata/1/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/prodata/1/iaa/6CTUS\_COMB.pep:\*  
6: /cgn2\_6/prodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description       |
|------------|-------|-------------|--------|-------|-------------------|
| 1          | 171   | 100.0       | 31     | 3     | US-09-105-678A-9  |
| 2          | 171   | 100.0       | 31     | 3     | US-09-105-678A-43 |
| 3          | 171   | 100.0       | 31     | 3     | US-08-776-971-61  |
| 4          | 171   | 100.0       | 31     | 3     | US-09-421-208-9   |
| 5          | 171   | 100.0       | 31     | 3     | US-09-421-208-43  |
| 6          | 171   | 100.0       | 31     | 4     | US-09-560-915-15  |
| 7          | 171   | 100.0       | 32     | 3     | US-09-105-678A-44 |
| 8          | 171   | 100.0       | 32     | 3     | US-08-776-971-62  |
| 9          | 171   | 100.0       | 32     | 3     | US-09-421-208-44  |
| 10         | 171   | 100.0       | 33     | 3     | US-09-105-678A-45 |
| 11         | 171   | 100.0       | 33     | 3     | US-08-776-971-63  |
| 12         | 171   | 100.0       | 33     | 3     | US-09-421-208-45  |
| 13         | 171   | 100.0       | 87     | 3     | US-08-776-971-59  |
| 14         | 171   | 100.0       | 87     | 3     | US-08-776-971-135 |
| 15         | 171   | 100.0       | 87     | 3     | US-08-776-971-138 |
| 16         | 158   | 92.4        | 31     | 3     | US-09-105-678A-7  |
| 17         | 158   | 92.4        | 31     | 3     | US-09-105-678A-31 |
| 18         | 158   | 92.4        | 31     | 3     | US-08-776-971-5   |
| 19         | 158   | 92.4        | 31     | 3     | US-08-776-971-97  |
| 20         | 158   | 92.4        | 31     | 3     | US-09-421-208-7   |
| 21         | 158   | 92.4        | 31     | 3     | US-09-421-208-31  |
| 22         | 158   | 92.4        | 31     | 4     | US-09-560-915-13  |
| 23         | 158   | 92.4        | 32     | 3     | US-09-105-678A-32 |
| 24         | 158   | 92.4        | 32     | 3     | US-08-776-971-6   |
| 25         | 158   | 92.4        | 32     | 3     | US-09-421-208-32  |
| 26         | 158   | 92.4        | 33     | 3     | US-09-105-678A-33 |
| 27         | 158   | 92.4        | 33     | 3     | US-08-776-971-7   |

Sequence 33, Appli  
Sequence 1, Appli  
Sequence 44, Appli  
Sequence 122, App  
Sequence 131, App  
Sequence 115, App  
Sequence 117, App  
Sequence 8, Appli  
Sequence 37, Appli  
Sequence 4, Appli  
Sequence 47, Appli  
Sequence 8, Appli  
Sequence 37, Appli  
Sequence 14, Appli  
Sequence 4, Appli  
Sequence 38, Appli  
Sequence 48, Appli

28 158 92.4 33 3 US-09-421-208-33  
29 158 92.4 98 3 US-08-776-971-1  
30 158 92.4 98 3 US-08-776-971-44  
31 158 92.4 98 3 US-08-776-971-122  
32 158 92.4 98 3 US-08-776-971-131  
33 158 92.4 98 3 US-08-776-971-136  
34 154 90.1 98 3 US-08-776-971-115  
35 154 90.1 98 3 US-08-776-971-117  
36 149 87.1 31 3 US-09-105-678A-8  
37 149 87.1 31 3 US-09-105-678A-37  
38 149 87.1 31 3 US-09-172-353-4  
39 149 87.1 31 3 US-08-776-971-47  
40 149 87.1 31 3 US-09-421-208-8  
41 149 87.1 31 3 US-09-421-208-37  
42 149 87.1 31 4 US-09-560-915-14  
43 149 87.1 31 4 US-09-799-955-4  
44 149 87.1 32 3 US-09-105-678A-38  
45 149 87.1 32 3 US-08-776-971-48

#### ALIGNMENTS

RESULT 1  
US-09-105-678A-9  
; Sequence 9, Application US/09105678A  
; Patent No. 6103882  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSES: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 31 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-9

Query Match 100.0%; Score 171; DB 3; Length 31;  
Best Local Similarity 100.0%; Pred. No. 7.3e-19;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRTSHSMEIRTPDINPAWYASRGIRPVGRF 31



COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/421,208  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/105,678  
FILING DATE: 26-JUN-1998  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 31 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-421-208-9

Query Match 100.0%; Score 171; DB 3; Length 31;  
Best Local Similarity 100.0%; Pred. No. 7.3e-19;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRTRHSMETPTDINPAWASRGIRPVGRF 31  
DB 1 SRTRHSMETPTDINPAWASRGIRPVGRF 31

RESULT 5  
US-09-421-208-43  
Sequence 43, Application US/09421208  
Patent No. 6258561  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/421,208  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/105,678  
FILING DATE: 26-JUN-1998  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 43:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 31 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-421-208-43

Query Match 100.0%; Score 171; DB 3; Length 31;  
Best Local Similarity 100.0%; Pred. No. 7.3e-19;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRTRHSMETPTDINPAWASRGIRPVGRF 31  
DB 1 SRTRHSMETPTDINPAWASRGIRPVGRF 31

RESULT 6  
US-09-560-915-15  
Sequence 15, Application US/09560915  
Patent No. 6383764  
GENERAL INFORMATION:  
APPLICANT: Lin, Steven  
APPLICANT: Civelli, Olivier  
TITLE OF INVENTION: Therapeutic Compositions and Methods  
RELATING TO Prolactin Releasing Peptide (PrRP)  
FILE REFERENCE: P-UC 3534  
CURRENT APPLICATION NUMBER: US/09/560,915  
CURRENT FILING DATE: 2000-04-28  
NUMBER OF SEQ ID NOS: 24  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 15  
LENGTH: 31  
TYPE: Prt  
ORGANISM: Homo Sapien  
US-09-560-915-15

Query Match 100.0%; Score 171; DB 4; Length 31;  
Best Local Similarity 100.0%; Pred. No. 7.3e-19;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRTRHSMETPTDINPAWASRGIRPVGRF 31  
DB 1 SRTRHSMETPTDINPAWASRGIRPVGRF 31

RESULT 7  
US-09-105-678A-44  
Sequence 44, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A

;; FILING DATE: 26-JUN-1998  
;; PRIOR APPLICATION DATA: JP 172118/1997  
;; FILING DATE: 27-JUN-1997  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Conlin, David G.  
;; REGISTRATION NUMBER: 27,026  
;; REFERENCE/DOCKET NUMBER: 48466-342  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 617-523-3400  
;; TELEFAX: 617-523-6440  
;; INFORMATION FOR SEQ ID NO: 44:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 32 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: linear  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-09-105-678A-44

Query Match 100.0%; Score 171; DB 3; Length 32;  
Best Local Similarity 100.0%; Pred. No. 7.6e-19;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRTRHSMETPDINPAMYASRGIRPVGRF 31  
Db 1 SRTRHSMETPDINPAMYASRGIRPVGRF 31

RESULT 8  
US-08-776-971-62  
; Sequence 62, Application US/08776971B  
; Patent No. 6228984  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; Habata, Yuco  
; Kawamata, Yuji  
; Hosoya, Masaki  
; Fujii, Ryo  
; Fukusumi, Shoji  
; Kitada, Chieko  
; TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
; NUMBER OF SEQUENCES: 140  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Fast-SEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/776,971B  
; FILING DATE: 06-Feb-1997  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/JP96/03821  
; FILING DATE: 28-DEC-1996  
; APPLICATION NUMBER: JP 7/343371  
; FILING DATE: 28-DEC-1995  
; APPLICATION NUMBER: JP 8/59419  
; FILING DATE: 15-MAR-1996  
; APPLICATION NUMBER: JP 8/211805  
; FILING DATE: 12-AUG-1996  
; APPLICATION NUMBER: JP 8/246573  
; FILING DATE: 18-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026

;; REFERENCE/DOCKET NUMBER: 47176  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 617-523-3400  
;; TELEFAX: 617-523-6440  
;; INFORMATION FOR SEQ ID NO: 62:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 32 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; FRAGMENT TYPE: internal  
;; SEQUENCE DESCRIPTION: SEQ ID NO: 62:  
US-08-776-971-62

Query Match 100.0%; Score 171; DB 3; Length 32;  
Best Local Similarity 100.0%; Pred. No. 7.6e-19;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRTRHSMETPDINPAMYASRGIRPVGRF 31  
Db 1 SRTRHSMETPDINPAMYASRGIRPVGRF 31

RESULT 9  
US-09-421-208-44  
; Sequence 44, Application US/09421208  
; Patent No. 6258561  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/421,208  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/105,678  
; FILING DATE: 26-JUN-1998  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 44:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 32 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-421-208-44

Query Match 100.0%; Score 171; DB 3; Length 32;  
Best Local Similarity 100.0%; Pred. No. 7.6e-19;

Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRTHRSMEIRTPDINPAWYASRGIRPVGRF 31  
 Db 1 SRTHRSMEIRTPDINPAWYASRGIRPVGRF 31

RESULT 10  
 US-09-105-678A-45  
 ; Sequence 45, Application US/09105678A  
 ; Patent No. 6103882  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Suenaga, Masato  
 ; APPLICANT: Moriya, Takeo  
 ; APPLICANT: Tanaka, Yoko  
 ; APPLICANT: Nishimura, Osamu  
 ; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
 ; NUMBER OF SEQUENCES: 52  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
 ; STREET: 130 Water Street  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02109

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/105,678A  
 FILING DATE: 26-JUN-1998

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 172118/1997

FILING DATE: 27-JUN-1997  
 ATTORNEY/AGENT INFORMATION:

NAME: Conlin, David G.  
 REGISTRATION NUMBER: 27,026

REFERENCE/DOCKET NUMBER: 48466-342  
 TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-523-3400  
 TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 45:

SEQUENCE CHARACTERISTICS:

LENGTH: 33 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-09-105-678A-45

Query Match 100.0%; Score 171; DB 3; Length 33;  
 Best Local Similarity 100.0%; Pred. No. 7.9e-19;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRTHRSMEIRTPDINPAWYASRGIRPVGRF 31  
 Db 1 SRTHRSMEIRTPDINPAWYASRGIRPVGRF 31

RESULT 11  
 US-08-776-971-63  
 ; Sequence 63, Application US/08776971B  
 ; Patent No. 6228984  
 ; GENERAL INFORMATION:

APPLICANT: Hinuma, Shuji  
 Habata, Yugo  
 Kawamata, Yuji  
 Hosoya, Masaki  
 Fujii, Ryo  
 Fukusumi, Shoji  
 Kitada, Chieko

; TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
 ; NUMBER OF SEQUENCES: 140  
 ; CORRESPONDENCE ADDRESS:

ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
 STREET: 130 Water Street  
 CITY: Boston  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/776,971B

FILING DATE: 06-Feb-1997

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/JP96/03821

FILING DATE: 28-DEC-1996

APPLICATION NUMBER: JP 7/343371

FILING DATE: 28-DEC-1995

APPLICATION NUMBER: JP 8/59419

FILING DATE: 15-MAR-1996

APPLICATION NUMBER: JP 8/211805

FILING DATE: 12-AUG-1996

APPLICATION NUMBER: JP 8/246573

FILING DATE: 18-SEP-1996

ATTORNEY/AGENT INFORMATION:

NAME: Conlin, David G.

REGISTRATION NUMBER: 27,026

REFERENCE/DOCKET NUMBER: 47176

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 63:

SEQUENCE CHARACTERISTICS:

LENGTH: 33 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

FRAGMENT TYPE: internal

SEQUENCE DESCRIPTION: SEQ ID NO: 63:

US-08-776-971-63

Query Match 100.0%; Score 171; DB 3; Length 33;  
 Best Local Similarity 100.0%; Pred. No. 7.9e-19;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRTHRSMEIRTPDINPAWYASRGIRPVGRF 31  
 Db 1 SRTHRSMEIRTPDINPAWYASRGIRPVGRF 31

RESULT 12  
 US-09-421-208-45  
 ; Sequence 45, Application US/09421208  
 ; Patent No. 6258561  
 ; GENERAL INFORMATION:

APPLICANT: Suenaga, Masato  
 Moriya, Takeo  
 Tanaka, Yoko  
 Nishimura, Osamu  
 TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
 NUMBER OF SEQUENCES: 52  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
 STREET: 130 Water Street  
 CITY: Boston  
 STATE: MA  
 COUNTRY: USA

ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/421,208  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/105,678  
FILING DATE: 26-JUN-1998  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 45:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 33 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-421-208-45

Query Match 100.0%; Score 171; DB 3; Length 33;  
Best Local Similarity 100.0%; Pred. No. 7.9e-19;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRTRHSMEIRTPDINPAWYASRGIRPVGRF 31  
Db 1 SRTRHSMEIRTPDINPAWYASRGIRPVGRF 31

RESULT 13  
US-08-776-971-59  
Sequence 59, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fujii, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371

FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 59:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 87 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 59:  
US-08-776-971-59

Query Match 100.0%; Score 171; DB 3; Length 87;  
Best Local Similarity 100.0%; Pred. No. 2.4e-18;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRTRHSMEIRTPDINPAWYASRGIRPVGRF 31  
Db 23 SRTRHSMEIRTPDINPAWYASRGIRPVGRF 53

RESULT 14  
US-08-776-971-135  
Sequence 135, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fujii, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996

```
; APPLICATION NUMBER: JP 8/246573
; FILING DATE: 18-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 47176
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 135:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 87 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 135:
US-08-776-971-135

Query Match      100.0%; Score 171; DB 3; Length 87;
Best Local Similarity 100.0%; Pred. No. 2.4e-18;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 SRTHSHMEIRTPDINPAWYASRGIRPVGRP 31
Db      23 SRTHSHMEIRTPDINPAWYASRGIRPVGRP 53

; TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
; NUMBER OF SEQUENCES: 140
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/776,971B
; FILING DATE: 06-Feb-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP96/03821
; FILING DATE: 28-DEC-1996
; APPLICATION NUMBER: JP 7/343371
; FILING DATE: 28-DEC-1995
; APPLICATION NUMBER: JP 8/59419
; FILING DATE: 15-MAR-1996
; APPLICATION NUMBER: JP 8/211805
; FILING DATE: 12-AUG-1996
; APPLICATION NUMBER: JP 8/246573
; FILING DATE: 18-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
```

```
; REFERENCE/DOCKET NUMBER: 47176
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 138:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 87 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 138:
US-08-776-971-138

Query Match      100.0%; Score 171; DB 3; Length 87;
Best Local Similarity 100.0%; Pred. No. 2.4e-18;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 SRTHSHMEIRTPDINPAWYASRGIRPVGRP 31
Db      23 SRTHSHMEIRTPDINPAWYASRGIRPVGRP 53

Search completed: December 3, 2003, 19:15:55
Job time : 11.5 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: December 3, 2003, 19:11:30 ; Search time 22.5 Seconds  
(without alignments)  
256.244 Million cell updates/sec

Title: US-09-868-885B-32  
Perfect score: 171  
Sequence: 1 SRTRHSMEIRTPDINPAWASRGIRPVGRF 31

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database : Published Applications AA:\*

```
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US03_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description      |
|------------|-------|-------------|--------|-------|------------------|
| 1          | 171   | 100.0       | 31     | 9     | US-09-932-161-15 |
| 2          | 171   | 100.0       | 31     | 12    | US-10-096-777-15 |
| 3          | 171   | 100.0       | 87     | 14    | US-10-044-592-92 |
| 4          | 158   | 92.4        | 31     | 9     | US-09-932-161-13 |
| 5          | 158   | 92.4        | 31     | 12    | US-10-096-777-13 |
| 6          | 158   | 92.4        | 31     | 14    | US-10-044-592-39 |
| 7          | 158   | 92.4        | 32     | 14    | US-10-044-592-40 |
| 8          | 158   | 92.4        | 33     | 14    | US-10-044-592-41 |
| 9          | 158   | 92.4        | 98     | 14    | US-10-044-592-28 |
| 10         | 158   | 92.4        | 98     | 14    | US-10-044-592-38 |
| 11         | 158   | 92.4        | 98     | 14    | US-10-044-592-82 |
| 12         | 158   | 92.4        | 98     | 14    | US-10-044-592-94 |
| 13         | 158   | 92.4        | 98     | 14    | US-10-044-592-86 |
| 14         | 158   | 92.4        | 98     | 14    | US-10-044-592-88 |
| 15         | 149   | 87.1        | 31     | 9     | US-09-932-161-14 |

|    |     |      |     |    |                     |                    |
|----|-----|------|-----|----|---------------------|--------------------|
| 16 | 149 | 87.1 | 31  | 12 | US-10-096-777-14    | Sequence 14, Appl  |
| 17 | 149 | 87.1 | 31  | 14 | US-10-044-592-4     | Sequence 4, Appl   |
| 18 | 149 | 87.1 | 31  | 14 | US-10-044-592-5     | Sequence 5, Appl   |
| 19 | 149 | 87.1 | 70  | 14 | US-10-044-592-90    | Sequence 90, Appl  |
| 20 | 149 | 87.1 | 82  | 14 | US-10-044-592-1     | Sequence 1, Appl   |
| 21 | 149 | 87.1 | 86  | 14 | US-10-044-592-96    | Sequence 96, Appl  |
| 22 | 149 | 87.1 | 91  | 14 | US-10-044-592-94    | Sequence 94, Appl  |
| 23 | 147 | 86.0 | 29  | 14 | US-10-044-592-26    | Sequence 26, Appl  |
| 24 | 128 | 74.9 | 25  | 14 | US-10-044-592-78    | Sequence 78, Appl  |
| 25 | 113 | 66.1 | 20  | 9  | US-09-932-161-18    | Sequence 18, Appl  |
| 26 | 113 | 66.1 | 20  | 12 | US-10-096-777-18    | Sequence 18, Appl  |
| 27 | 109 | 63.7 | 20  | 9  | US-09-932-161-16    | Sequence 16, Appl  |
| 28 | 109 | 63.7 | 20  | 12 | US-10-096-777-16    | Sequence 16, Appl  |
| 29 | 109 | 63.7 | 20  | 14 | US-10-044-592-42    | Sequence 42, Appl  |
| 30 | 109 | 63.7 | 21  | 14 | US-10-044-592-43    | Sequence 43, Appl  |
| 31 | 109 | 63.7 | 22  | 14 | US-10-044-592-44    | Sequence 44, Appl  |
| 32 | 105 | 61.4 | 20  | 9  | US-09-932-161-17    | Sequence 17, Appl  |
| 33 | 105 | 61.4 | 20  | 12 | US-10-096-777-17    | Sequence 17, Appl  |
| 34 | 105 | 61.4 | 20  | 14 | US-10-044-592-6     | Sequence 6, Appl   |
| 35 | 103 | 60.2 | 19  | 14 | US-10-044-592-27    | Sequence 27, Appl  |
| 36 | 88  | 51.5 | 40  | 14 | US-10-044-592-80    | Sequence 80, Appl  |
| 37 | 57  | 33.3 | 9   | 14 | US-10-044-592-8     | Sequence 8, Appl   |
| 38 | 54  | 31.6 | 209 | 14 | US-10-108-915-30    | Sequence 30, Appl  |
| 39 | 54  | 31.6 | 428 | 9  | US-09-820-155-4     | Sequence 4, Appl   |
| 40 | 51  | 29.8 | 465 | 12 | US-10-301-822-197   | Sequence 197, Appl |
| 41 | 50  | 29.2 | 428 | 9  | US-09-820-155-2     | Sequence 2, Appl   |
| 42 | 49  | 28.7 | 402 | 15 | US-10-156-761-10748 | Sequence 10748, A  |
| 43 | 48  | 28.1 | 10  | 14 | US-10-044-592-9     | Sequence 9, Appl   |
| 44 | 48  | 28.1 | 159 | 15 | US-10-106-698-7900  | Sequence 7900, Ap  |
| 45 | 48  | 28.1 | 171 | 15 | US-10-187-267A-5    | Sequence 5, Appl   |

## ALIGNMENTS

### RESULT 1

US-09-932-161-15  
; Sequence 15, Application US/09932161  
; Patent No. US20020037533A1  
; GENERAL INFORMATION:  
; APPLICANT: Civealli, Olivier  
; APPLICANT: Lin, Steven  
; TITLE OF INVENTION: Screening and Therapeutic Methods For  
; TITLE OF INVENTION: Promoting Wakefulness and Sleep  
; FILE REFERENCE: P-UC 4679  
; CURRENT APPLICATION NUMBER: US/09/932.161  
; PRIOR FILING DATE: 2001-08-17  
; PRIOR APPLICATION NUMBER: US 09/560,915  
; PRIOR FILING DATE: 2000-04-28  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15  
; LENGTH: 31  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-09-932-161-15

Query Match 100.0%; Score 171; DB 9; Length 31;  
Best Local Similarity 100.0%; Pred. No. 1.1e-17;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRTRHSMEIRTPDINPAWASRGIRPVGRF 31  
DB 1 SRTRHSMEIRTPDINPAWASRGIRPVGRF 31

### RESULT 2

US-10-096-777-15  
; Sequence 15, Application US/10096777  
; Publication No. US20030171270A1  
; GENERAL INFORMATION:  
; APPLICANT: Civealli, Olivier  
; APPLICANT: Lin, Steven

```

Query Match      92.4%; Score 158; DB 9; Length 31;
Best Local Similarity 90.3%; Pred. No. 8.5e-16;
Matches 28; Conservative 1; Mismatches -2; Indels

QY      1 SRTSHSMEIRTPDINPAWYASRGIRPVGRE 31
      ||| : ||| ||| ||| ||| ||| ||| ||| |||
DB      1 SRAHSHMEIRTPDINPAWYASRGIRPVGRE 31
      ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 5  
US-10-096-777-13  
; Sequence 13, Application US/10096777  
; Publication No. US20030171270A1

```

RESULT 5
US-10-096-777-13
; Sequence 13, Application US/10096777
; Publication NO. US20030171270A1
; GENERAL INFORMATION:
; APPLICANT: Civelli, Olivier
; APPLICANT: Lin, Steven
; TITLE OF INVENTION: Therapeutic Compositions and Methods
; TITLE OF INVENTION: Relating To Prolactin Releasing Peptide (PrRP)
; FILE REFERENCE: P-UC 3534

```

```
Query Match      92.4%; Score 158; DB 12; Length 31;  
Best Local Similarity 90.3%; Pred. No. 8.5e-16;  
Matches 28; Conservative 1; Mismatches -2; Indels  
  
QY    1   SRTRHMEIRTPDINPAWASGRIPVGRF 31  
      |||||  
Db     1   SRAOHSEIRTPDINPAWAGRGIRPVGRF 31
```

```

RESULT 6
US-10-044-592-39
; Sequence 39, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Fukusumi, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96

```

```

RESULT 6
US-10-044-592-39
/ Sequence 39, Application US/10044592
/ Publication No. US20020143152A1
/ GENERAL INFORMATION:
/ APPLICANT: Hinuma, Shuji
/ APPLICANT: Fukusumi Shoji
/ TITLE OF INVENTION: Polypeptides, their Production and Use
/ FILE OF INVENTION: Polypeptides
/ CURRENT APPLICATION NUMBER: US/10/044,592
/ CURRENT FILING DATE: 2002-01-10
/ PRIOR APPLICATION NUMBER: US 09/403639
/ PRIOR FILING DATE: 1999-25-10
/ PRIOR APPLICATION NUMBER: PCT/JP98/01923
/ PRIOR FILING DATE: 1998-04-27
/ PRIOR APPLICATION NUMBER: JP 9-109974
/ PRIOR FILING DATE: 1997-04-28
/ NUMBER OF SEQ ID NOS: 96
/ SOFTWARE:
/ SEQ ID NO 39
/ LENGTH: 31
/ TYPE: PRT
/ ORGANISM: Bovine
US-10-044-592-39

```

Query Match 92.4%; Score 158; DB 14; Length 31;

Best Local Similarity 90.3%; Pred. No. 8.5e-16; DB 14; Length 32;  
Matches 28; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SRTRHSMETPTDINPAWYASRGIRPVGRF 31  
||:|||||||||||||||||||||||  
Db 1 SRAHQHSMETPTDINPAWYAGRGIRPVGRF 31  
||:|||||||||||||||||||||||

RESULT 7  
US-10-044-592-40  
; Sequence 40, Application US/10044592  
; Publication No. US20020143152A1  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; TITLE OF INVENTION: Polypeptides, their Production and Use  
; FILE REFERENCE: 2463US2P  
; CURRENT APPLICATION NUMBER: US/10/044,592  
; CURRENT FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: US 09/403639  
; PRIOR FILING DATE: 1999-25-10  
; PRIOR APPLICATION NUMBER: PCT/JP98/01923  
; PRIOR FILING DATE: 1998-04-27  
; PRIOR APPLICATION NUMBER: JP 9-109974  
; PRIOR FILING DATE: 1997-04-28  
; NUMBER OF SEQ ID NOS: 96  
; SOFTWARE:  
; SEQ ID NO 40  
; LENGTH: 32  
; TYPE: PRT  
; ORGANISM: Bovine  
US-10-044-592-40

Query Match 92.4%; Score 158; DB 14; Length 32;  
Best Local Similarity 90.3%; Pred. No. 8.8e-16;  
Matches 28; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SRTRHSMETPTDINPAWYASRGIRPVGRF 31  
||:|||||||||||||||||||||||  
Db 1 SRAHQHSMETPTDINPAWYAGRGIRPVGRF 31  
||:|||||||||||||||||||||||

RESULT 8  
US-10-044-592-41  
; Sequence 41, Application US/10044592  
; Publication No. US20020143152A1  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; TITLE OF INVENTION: Polypeptides, their Production and Use  
; FILE REFERENCE: 2463US2P  
; CURRENT APPLICATION NUMBER: US/10/044,592  
; CURRENT FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: US 09/403639  
; PRIOR FILING DATE: 1999-25-10  
; PRIOR APPLICATION NUMBER: PCT/JP98/01923  
; PRIOR FILING DATE: 1998-04-27  
; PRIOR APPLICATION NUMBER: JP 9-109974  
; PRIOR FILING DATE: 1997-04-28  
; NUMBER OF SEQ ID NOS: 96  
; SOFTWARE:  
; SEQ ID NO 41  
; LENGTH: 33  
; TYPE: PRT  
; ORGANISM: Bovine  
US-10-044-592-41

Query Match 92.4%; Score 158; DB 14; Length 33;  
Best Local Similarity 90.3%; Pred. No. 9.1e-16;  
Matches 28; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SRTRHSMETPTDINPAWYASRGIRPVGRF 31  
||:|||||||||||||||||||||||

Db 1 SRAHQHSMETPTDINPAWYAGRGIRPVGRF 31

RESULT 9  
US-10-044-592-28  
; Sequence 28, Application US/10044592  
; Publication No. US20020143152A1  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; TITLE OF INVENTION: Polypeptides, their Production and Use  
; FILE REFERENCE: 2463US2P  
; CURRENT APPLICATION NUMBER: US/10/044,592  
; CURRENT FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: US 09/403639  
; PRIOR FILING DATE: 1999-25-10  
; PRIOR APPLICATION NUMBER: PCT/JP98/01923  
; PRIOR FILING DATE: 1998-04-27  
; PRIOR APPLICATION NUMBER: JP 9-109974  
; PRIOR FILING DATE: 1997-04-28  
; NUMBER OF SEQ ID NOS: 96  
; SOFTWARE:  
; SEQ ID NO 28  
; LENGTH: 98  
; TYPE: PRT  
; ORGANISM: Murine  
US-10-044-592-28

Query Match 92.4%; Score 158; DB 14; Length 98;  
Best Local Similarity 90.3%; Pred. No. 2.8e-15;  
Matches 28; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SRTRHSMETPTDINPAWYASRGIRPVGRF 31  
||:|||||||||||||||||||||||  
Db 23 SRAHQHSMETPTDINPAWYAGRGIRPVGRF 53  
||:|||||||||||||||||||||||

RESULT 10  
US-10-044-592-38  
; Sequence 38, Application US/10044592  
; Publication No. US20020143152A1  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; TITLE OF INVENTION: Polypeptides, their Production and Use  
; FILE REFERENCE: 2463US2P  
; CURRENT APPLICATION NUMBER: US/10/044,592  
; CURRENT FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: US 09/403639  
; PRIOR FILING DATE: 1999-25-10  
; PRIOR APPLICATION NUMBER: PCT/JP98/01923  
; PRIOR FILING DATE: 1998-04-27  
; PRIOR APPLICATION NUMBER: JP 9-109974  
; PRIOR FILING DATE: 1997-04-28  
; NUMBER OF SEQ ID NOS: 96  
; SOFTWARE:  
; SEQ ID NO 38  
; LENGTH: 98  
; TYPE: PRT  
; ORGANISM: Bovine  
US-10-044-592-38

Query Match 92.4%; Score 158; DB 14; Length 98;  
Best Local Similarity 90.3%; Pred. No. 2.8e-15;  
Matches 28; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SRTRHSMETPTDINPAWYASRGIRPVGRF 31  
||:|||||||||||||||||||||||  
Db 23 SRAHQHSMETPTDINPAWYAGRGIRPVGRF 53  
||:|||||||||||||||||||||||

RESULT 11  
US-10-044-592-82

```
; Sequence 82, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 82
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Bovine
; US-10-044-592-82

Query Match          92.4%; Score 158; DB 14; Length 98;
Best Local Similarity 90.3%; Pred. No. 2.8e-15;
Matches 28; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SRTSHSMEIRTPDINPAWYASRGIRPVGRF 31
Db 23 SRAHQHSMIEIRTPDINPAWYAGRGIRPVGRF 53

RESULT 12
US-10-044-592-84
; Sequence 84, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 84
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Bovine
; US-10-044-592-84

Query Match          92.4%; Score 158; DB 14; Length 98;
Best Local Similarity 90.3%; Pred. No. 2.8e-15;
Matches 28; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SRTSHSMEIRTPDINPAWYASRGIRPVGRF 31
Db 23 SRAHQHSMIEIRTPDINPAWYAGRGIRPVGRF 53

RESULT 13
US-10-044-592-86
; Sequence 86, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Fukusumi, Shoji
```

```
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 86
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Bovine
; US-10-044-592-86

Query Match          92.4%; Score 158; DB 14; Length 98;
Best Local Similarity 90.3%; Pred. No. 2.8e-15;
Matches 28; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SRTSHSMEIRTPDINPAWYASRGIRPVGRF 31
Db 23 SRAHQHSMIEIRTPDINPAWYAGRGIRPVGRF 53

RESULT 14
US-10-044-592-88
; Sequence 88, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Fukusumi, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 88
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Bovine
; US-10-044-592-88

Query Match          92.4%; Score 158; DB 14; Length 98;
Best Local Similarity 90.3%; Pred. No. 2.8e-15;
Matches 28; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SRTSHSMEIRTPDINPAWYASRGIRPVGRF 31
Db 23 SRAHQHSMIEIRTPDINPAWYAGRGIRPVGRF 53

RESULT 15
US-09-932-161-14
; Sequence 14, Application US/09932161
; Patent No. US20020037533A1
; GENERAL INFORMATION:
; APPLICANT: Civeilli, Olivier
; APPLICANT: Lin, Steven
; TITLE OF INVENTION: Screening and Therapeutic Methods For
; FILE REFERENCE: P-UC 4679
; CURRENT APPLICATION NUMBER: US/09/932,161
; CURRENT FILING DATE: 2001-08-17
```

; PRIOR APPLICATION NUMBER: US 09/560,915  
; PRIOR FILING DATE: 2000-04-28  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14  
; LENGTH: 31  
; TYPE: PRT  
; ORGANISM: Rattus  
US-09-932-161-14

Query Match 87.1%; Score 149; DB 9; Length 31;  
Best Local Similarity 83.9%; Pred. No. 1.7e-14;  
Matches 26; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 SRTHRSMEITPDINPAWYASRGIRPVGRF 31  
|||:|||||  
Db 1 SRAHQSMETRTPDINPAWYTGIRPVGRF 31  
|||:|||||

Search completed: December 3, 2003, 19:17:37  
Job time : 22.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 19:05:30 ; Search time 10.5 Seconds  
(without alignments)  
283.927 Million cell updates/sec

Title: US-09-868-885B-32

Perfect score: 171

Sequence: 1 SRTRHSMEIRTPDINPAWYASRGIRPVGRF 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Pir1.\*  
2: Pir2.\*  
3: Pir3.\*  
4: Pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Match | Query Length | ID       | Description        |
|------------|-------|-------|--------------|----------|--------------------|
| 1          | 149   | 87.1  | 83           | 2 JC7607 | prolactin-releasin |
| 2          | 56.5  | 33.0  | 664          | 2 F83376 | conserved hypothet |
| 3          | 54    | 31.6  | 790          | 2 T47959 | hypothetical prote |
| 4          | 53.5  | 31.3  | 314          | 2 B70569 | hypothetical prote |
| 5          | 53    | 31.0  | 637          | 2 D87559 | sensory box histid |
| 6          | 52    | 30.4  | 785          | 2 F89099 | sensory transducti |
| 7          | 50.5  | 29.5  | 1882         | 2 S73484 | hypothetical prote |
| 8          | 50    | 29.2  | 123          | 2 S77900 | hypothetical prote |
| 9          | 50    | 29.2  | 128          | 2 S76955 | hypothetical prote |
| 10         | 49.5  | 28.9  | 176          | 2 S67150 | hypothetical prote |
| 11         | 49    | 28.7  | 72           | 2 B91002 | probable regulator |
| 12         | 48.5  | 28.4  | 303          | 2 A82016 | hypothetical prote |
| 13         | 48.5  | 28.4  | 1501         | 2 T45623 | hypothetical prote |
| 14         | 48    | 28.1  | 118          | 2 AC3169 | hypothetical prote |
| 15         | 48    | 28.1  | 220          | 2 C83292 | probable glutathio |
| 16         | 48    | 28.1  | 335          | 2 S70671 | lipopolysaccharide |
| 17         | 48    | 28.1  | 348          | 2 T21648 | hypothetical prote |
| 18         | 48    | 28.1  | 455          | 2 D70885 | probable alcd prot |
| 19         | 48    | 28.1  | 1084         | 2 T33759 | hypothetical prote |
| 20         | 47.5  | 27.8  | 345          | 2 D84012 | N-acetylglutamate  |
| 21         | 47.5  | 27.8  | 393          | 2 A82664 | conserved hypothet |
| 22         | 47.5  | 27.8  | 401          | 2 A97446 | hypothetical prote |
| 23         | 47.5  | 27.8  | 503          | 2 A82193 | Sun/nucleolar prot |
| 24         | 47.5  | 27.8  | 533          | 2 A33111 | segmentation prote |
| 25         | 47    | 27.5  | 159          | 2 G82669 | ubiquinone biosynt |
| 26         | 47    | 27.5  | 215          | 2 B87577 | glutathione S-tran |
| 27         | 47    | 27.5  | 284          | 2 F71015 | hypothetical prote |
| 28         | 47    | 27.5  | 333          | 2 H82852 | hydroxybenzoate oc |
| 29         | 47    | 27.5  | 501          | 2 T48336 | hypothetical prote |

|    |      |      |      |          |                    |
|----|------|------|------|----------|--------------------|
| 30 | 47   | 27.5 | 938  | 2 C84480 | hypothetical prote |
| 31 | 47   | 27.5 | 4589 | 2 T14914 | dynein beta heavy  |
| 32 | 46.5 | 27.2 | 240  | 2 D64688 | probable 1-acylgly |
| 33 | 46.5 | 27.2 | 779  | 2 T49717 | related to BCS1 pr |
| 34 | 46.5 | 27.2 | 957  | 2 A84089 | hypothetical prote |
| 35 | 46.5 | 27.2 | 1495 | 2 T31434 | densin-180 - rat   |
| 36 | 46.5 | 27.2 | 1607 | 2 T13250 | hypothetical prote |
| 37 | 46.5 | 27.2 | 1693 | 1 MNMWHB | genome polyprotein |
| 38 | 46.5 | 27.2 | 4957 | 2 T03455 | ALR protein - huma |
| 39 | 46.5 | 27.2 | 5262 | 2 T03454 | ALR protein - huma |
| 40 | 46   | 26.9 | 256  | 2 F70812 | probable lpqr prot |
| 41 | 46   | 26.9 | 342  | 2 B64395 | malic acid transpo |
| 42 | 46   | 26.9 | 347  | 2 H64371 | malic acid transpo |
| 43 | 46   | 26.9 | 419  | 2 A83166 | hypothetical prote |
| 44 | 46   | 26.9 | 688  | 2 A12516 | hypothetical prote |
| 45 | 46   | 26.9 | 698  | 2 T39050 | hypothetical prote |

#### ALIGNMENTS

##### RESULT 1

JC7607

prolactin-releasing peptide - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 30-Jun-2001

C;Accession: JC7607

K;Yamada, M.; Ozawa, A.; Ishii, S.; Shibusawa, N.; Hashida, T.; Hosoya, T

Biochem. Biophys. Res. Commun. 281, 53-56, 2001

A;Title: Isolation and characterization of the rat prolactin-releasing peptide gene: Mu

A;Reference number: JC7607; MUID:21092785; PMID:11178959

A;Contents: Spleen

A;Accession: JC7607

A;Molecule type: DNA

A;Residues: 1-83 <YAM>

A;Cross-references: DDBJ:AB040612; DDBJ:AB040613

C;Comment: This peptide induces arachidonic acid metabolite release from rat anterior p

release, and stimulation of ACTH secretion from the pituitary.

C;Genetics:

A;Gene: PrRP

A;Introns: 33/1

Query Match 87.1%; Score 149; DB 2; Length 83;  
Best Local Similarity 83.9%; Pred. No. 8.7e-15;  
Matches 26; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 SRTRHSMEIRTPDINPAWYASRGIRPVGRF 31

Db 22 SRTRHSMEIRTPDINPAWYASRGIRPVGRF 52

##### RESULT 2

F83376

conserved hypothetical protein PA2151 [imported] - Pseudomonas aeruginosa (strain PA01)

C;Species: Pseudomonas aeruginosa

C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000

C;Accession: F83376

R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim

; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path

A;Reference number: A82950; MUID:20437337; PMID:10984043

A;Accession: F83376

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-664 <STO>

A;Cross-references: GB:AE004642; GB:AE004091; NID:g9948163; PIDN:AAG05539.1; GSPDB:GN00

A;Experimental source: strain PA01

C;Genetics:

A;Gene: PA2151

Query Match 33.0%; Score 56.5; DB 2; Length 664;

C;Species: Caulobacter crescentus  
C;Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
C;Accession: D87559  
R;Nierman, W.C.; Feidblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.; B. J.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolb, J.; Esmolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.N. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A;Title: Complete Genome Sequence of Caulobacter crescentus.  
A;Reference number: A87249; MUID:21173698; PMID:11259647  
A;Accession: D87559  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-637 <STO>  
A;Cross-references: GB:AE005673; NID:g13424056; PIDN:AAK24472.1; GSPDB:GN00148  
C;Genetics:  
A;Gene: CC2501

Query Match 31.0%; Score 53; DB 2; Length 637;  
Best Local Similarity 48.0%; Pred. No. 11;  
Matches 12; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 4 HRHSMBIRTPDINPAWYASRGIRPV 28  
|||:|||||:|:  
Db 22 HRDSDDLRSPAINFAAIRVRLRAV 46

RESULT 6  
F69099  
sensory transduction histidine kinase - Methanobacterium thermoautotrophicum (strain Dc  
C;Species: Methanobacterium thermoautotrophicum  
C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 22-Oct-1999  
C;Accession: F69099  
R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Gifu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, F.; S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N. J. Bacteriol. 179, 7135-7155, 1997  
A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: func  
A;Reference number: A69000; MUID:98037514; PMID:9371463  
A;Accession: F69099  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-785 <WTH>  
A;Cross-references: GB:AE000805; GB:AE000666; NID:g2621213; PIDN:AA84680.1; PID:g2621213  
A;Experimental source: strain Delta H  
C;Genetics:  
A;Gene: WTH174

Query Match 30.4%; Score 52; DB 2; Length 785;  
Best Local Similarity 46.4%; Pred. No. 19;  
Matches 13; Conservative 1; Mismatches 12; Indels 2; Gaps 1;

QY 4 HRHSMBIRTPDINPAW--YASRGIRPVG 29  
|||:|||||:|:  
Db 412 HRATFRIIRPDGNYRWVEYVDRPIRTDG 439

RESULT 7  
S73484  
hypothetical protein K05\_orf1882 - Mycoplasma pneumoniae (strain ATCC 29342)  
C;Species: Mycoplasma pneumoniae  
A;Variety: ATCC 29342  
C;Date: 27-Feb-1997 #sequence\_revision 25-Apr-1997 #text\_change 07-Dec-1999  
C;Accession: S73484; S62840  
R;Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkle, E.; Li, B.C.; Herrmann, R. Nucleic Acids Res. 24, 4420-4449, 1996  
A;Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae  
A;Reference number: S73327; MUID:97105885; PMID:8948633  
A;Accession: S73484  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-1882 <HIM>  
A;Cross-references: EMBL:AE000017; GB:U00089; NID:g1673812; PIDN:AA895806.1; PID:g1673812  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library November 1996

C;Accession: AH2016  
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,  
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yasuda, M.; Tabata, S.  
DNA Res. 8, 205-213, 2001  
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An.  
A;Reference number: AB1807; PMID:21595285; PMID:11759840



A:Accession: AH2016  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-303 <KUR>  
 A:Cross-references: GB:BA000019; PIDN:BA078052.1; PID:g17135506; GSPDB:GN00179  
 A:Experimental source: strain PCC 7120  
 C:Genetics:  
 A:Gene: all1666

Query Match 28.4%; Score 48.5; DB 2; Length 303;  
 Best Local Similarity 48.4%; Pred. No. 22;  
 Matches 15; Conservative 1; Mismatches 10; Indels 5; Gaps 2;

QY 4 HSHMEIRTPDINPAWY-----ASRGIRPVGR 30  
 DB 226 HEHSYE-RTRADTGTTLTCGAGAGNRPVGR 255

## RESULT 13

T45623  
 hypothetical protein F13G24.180 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 04-Feb-2000  
 C:Accession: T45623  
 R:Bevan, M.; Van Der Schueren, J.; Chuang, Y.J.; Voet, M.; Robben, J.; Volckaert, G.; Ba  
 submitted to the Protein Sequence Database, December 1999  
 A:Reference number: Z23009  
 A:Accession: T45623  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1501 <BEV>  
 A:Cross-references: EMBL:AL133421  
 A:Experimental source: cultivar Columbia; BAC clone F13G24  
 C:Genetics:  
 A:Map position: 5  
 A:Introns: 64/1; 739/3; 785/2; 1302/2; 1318/3; 1399/2; 1434/2  
 A:Note: F13G24.180

Query Match 28.4%; Score 48.5; DB 2; Length 1501;  
 Best Local Similarity 29.5%; Pred. No. 1.3e+02;  
 Matches 13; Conservative 7; Mismatches 9; Indels 15; Gaps 3;

QY 3 THHSMEIR-----TPDINPAWYASRG-----IRPV---GRF 31  
 DB 1169 TKHSASVGDHQQISQWAPSWYSQYGTFKNGLVQFVNDTGRF 1212

## RESULT 14

AC3169  
 hypothetical protein Atu5077 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)  
 C:Species: Agrobacterium tumefaciens  
 C>Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 18-Nov-2002  
 C:Accession: AC3169  
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I  
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell  
 ; Karp, P.; Romero, P.; Zhang, S.  
 Science 294, 2317-2323, 2001  
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
 ster, B.W.  
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
 A:Reference number: AB2577; MUID:21608550; PMID:11743193  
 A:Accession: AC3169  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-118 <KUR>  
 A:Cross-references: GB:AE008687; PIDN:AAL45769.1; PID:g17743503; GSPDB:GN00188  
 A:Experimental source: strain C58 (Dupont)  
 C:Genetics:  
 A:Gene: Atu5077  
 A:Genome: plasmid

Query Match 28.1%; Score 48; DB 2; Length 118;  
 Best Local Similarity 34.4%; Pred. No. 9.4;

Matches 11; Conservative 5; Mismatches 12; Indels 4; Gaps 1;  
 QY 2 RTHHSMEIRTPDI-----NPAWYASRGIRPVGR 29  
 DB 19 KLHKHNAARRRPDIGARVQVPATAPHRGLAGIG 50

## RESULT 15

CB3292  
 probable glutathione S-transferase PA2821 [imported] - Pseudomonas aeruginosa (strain I  
 C:Species: Pseudomonas aeruginosa  
 C>Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 23-Mar-2001  
 C:Accession: CB3292  
 R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; I  
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lin  
 ; Lory, S.; Olson, M.V.  
 Nature 406, 959-964, 2000  
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pat  
 A:Reference number: AB2950; MUID:20437337; PMID:10984043  
 A:Accession: CB3292  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-220 <STO>  
 A:Cross-references: GB:AE004709; GB:AE004091; NID:g9948904; PIDN:AAG06209.1; GSPDB:GN0  
 A:Experimental source: strain PA01  
 C:Genetics:  
 A:Gene: PA2821  
 C:Superfamily: plaiace glutathione transferase

Query Match 28.1%; Score 48; DB 2; Length 220;  
 Best Local Similarity 44.0%; Pred. No. 18;  
 Matches 11; Conservative 3; Mismatches 9; Indels 2; Gaps 1;

QY 6 HSMETRTDINPAWYASRGIRPVGR 30  
 DB 28 YQLEAIAFPQGPAMY--REISPLGR 50

Search completed: December 3, 2003, 19:14:56  
 Job time : 10.5 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 19:01:34 ; Search time 6.25 Seconds

(without alignments)  
233.252 Million cell updates/sec

Title: US-09-868-885b-32

Perfect score: 171

Sequence: 1 SRTTHSMEIRTPDINPAWASRGIRPVGRF 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID        | Description        |
|------------|-------|-------------|--------|--------------|--------------------|
| 1          | 171   | 100.0       | 87     | 1 PRRP_HUMAN | P81277 homo sapien |
| 2          | 158   | 92.4        | 98     | 1 PRRP_BOVIN | P81264 bos taurus  |
| 3          | 149   | 87.1        | 93     | 1 PRRP_RAT   | P81278 rattus norv |
| 4          | 54    | 31.6        | 428    | 1 NER3_HUMAN | Q9uq49 homo sapien |
| 5          | 52    | 30.4        | 288    | 1 Y587_PASMU | Q9cn56 pasteurilla |
| 6          | 50.5  | 25.5        | 1882   | 1 Y468_MYCPN | P75109 mycoplasma  |
| 7          | 50    | 25.2        | 428    | 1 NER3_BOVIN | Q97859 bos taurus  |
| 8          | 47.5  | 27.8        | 345    | 1 ARQC_BACHD | Q9k8v2 bacillus ha |
| 9          | 47.5  | 27.8        | 533    | 1 CNC_DROME  | P20482 drosophila  |
| 10         | 47    | 27.5        | 402    | 1 EX7L_STRCO | Q9fhn3 streptomyce |
| 11         | 47    | 27.5        | 576    | 1 EX11_HUMAN | Q92935 homo sapien |
| 12         | 46.5  | 27.2        | 240    | 1 PLSC_HELPY | O25903 helicobacte |
| 13         | 46.5  | 27.2        | 1693   | 1 POLN_HEVBU | P29324 hepatitis e |
| 14         | 46.5  | 27.2        | 1693   | 1 POLN_HEVMT | Q04610 hepatitis e |
| 15         | 46.5  | 27.2        | 1783   | 1 Y468_MYCGR | Q49460 mycoplasma  |
| 16         | 46    | 26.9        | 342    | 1 Y762_METJA | Q58172 methanococc |
| 17         | 46    | 26.9        | 347    | 1 Y576_METUA | Q57996 methanococc |
| 18         | 46    | 26.9        | 383    | 1 CYCR_CHRVI | O82947 chromatiu   |
| 19         | 46    | 26.9        | 417    | 1 EX7L_CORGL | Q8nrm3 corynebacte |
| 20         | 46    | 26.9        | 447    | 1 EX7L_CORF  | Q8tgp1 corynebacte |
| 21         | 46    | 26.9        | 1400   | 1 R1P1_SCHPO | Q96ud3 schizosacch |
| 22         | 45.5  | 26.6        | 239    | 1 FGEL_SINY3 | P74618 synechocyst |
| 23         | 45.5  | 26.6        | 407    | 1 Y116_MYCTU | P72052 mycobacteri |
| 24         | 45    | 26.3        | 213    | 1 SRN2_YEAST | Q99176 saccharomyc |
| 25         | 45    | 26.3        | 321    | 1 HTPX_AGRT5 | Q8ubm5 agrobacteri |
| 26         | 45    | 26.3        | 350    | 1 HNRO_DROME | P10181 drosophila  |
| 27         | 45    | 26.3        | 510    | 1 YCGB_SCOLI | P29013 escherichia |
| 28         | 45    | 26.3        | 798    | 1 UNR_RAT    | P18395 rattus norv |
| 29         | 45    | 26.3        | 940    | 1 UVRA_VIBCH | Q9kuw5 vibrio chol |
| 30         | 45    | 26.3        | 962    | 1 UVRA_METTH | O26543 methanobact |
| 31         | 45    | 26.3        | 973    | 1 UVRA_RHILO | Q98m36 rhizobium 1 |
| 32         | 45    | 26.3        | 973    | 1 UVRA_RHIME | P56899 rhizobium m |
| 33         | 44.5  | 26.0        | 860    | 1 VGI2_BPB03 | Q37893 bacterioph  |

#### ALIGNMENTS

##### RESULT 1

| ID | PRRP_HUMAN                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | STANDARD; | PRT; | 87 AA. |
|----|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------|------|--------|
| AC | P81277;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |           |      |        |
| DT | 30-MAY-2000 (Rel. 39, Created)                                                                                                                                                                                                                                                                                                                                                                                                                                                            |           |      |        |
| DT | 30-MAY-2000 (Rel. 39, Last sequence update)                                                                                                                                                                                                                                                                                                                                                                                                                                               |           |      |        |
| DT | 28-FEB-2003 (Rel. 41, Last annotation update)                                                                                                                                                                                                                                                                                                                                                                                                                                             |           |      |        |
| DE | Prolactin-releasing peptide precursor (PrRP) (Prolactin-releasing hormone) [Contains: Prolactin-releasing peptide PrRP31; Prolactin-releasing peptide PrRP20].                                                                                                                                                                                                                                                                                                                            |           |      |        |
| GN | PRH.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |           |      |        |
| OS | Homo sapiens (Human).                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |           |      |        |
| OC | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;                                                                                                                                                                                                                                                                                                                                                                                                                                   |           |      |        |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.                                                                                                                                                                                                                                                                                                                                                                                                                                |           |      |        |
| OX | NCBI_TaxID=9606;                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |           |      |        |
| EN | [1]                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |           |      |        |
| RP | SEQUENCE FROM N.A.                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |           |      |        |
| RC | TISSUE=Brain;                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |           |      |        |
| RX | MEDLINE=98268781; PubMed=9607765;                                                                                                                                                                                                                                                                                                                                                                                                                                                         |           |      |        |
| RA | Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S., Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M., Kurokawa T., Nishimura O., Onda H., Fujino M.;                                                                                                                                                                                                                                                                                                            |           |      |        |
| RT | "A prolactin-releasing peptide in the brain."                                                                                                                                                                                                                                                                                                                                                                                                                                             |           |      |        |
| RL | Nature 393:272-276(1998).                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |           |      |        |
| RN | [2]                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |           |      |        |
| RP | TISSUE SPECIFICITY.                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |           |      |        |
| RX | MEDLINE=99426652; PubMed=10498338;                                                                                                                                                                                                                                                                                                                                                                                                                                                        |           |      |        |
| RA | Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Habata Y., Hinuma S., Sekiguchi M., Kitada C., Kurokawa T., Nishimura O., Onda H., Fujino M., Sumino Y., Fujino M.;                                                                                                                                                                                                                                                                                                                        |           |      |        |
| RT | "Tissue distribution of prolactin-releasing peptide (PrRP) and its receptor."                                                                                                                                                                                                                                                                                                                                                                                                             |           |      |        |
| RL | Regul. Pept. 83:1-10(1999).                                                                                                                                                                                                                                                                                                                                                                                                                                                               |           |      |        |
| CC | -!- FUNCTION: Stimulates prolactin (PRL) release and regulates the expression of prolactin through its receptor GPR10. May stimulate lactotrophs directly to secrete PRL.                                                                                                                                                                                                                                                                                                                 |           |      |        |
| CC | -!- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.                                                                                                                                                                                                                                                                                                                                                                                                                               |           |      |        |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch). |           |      |        |
| CC | EMBL; AB015419; BAA29027.1; -                                                                                                                                                                                                                                                                                                                                                                                                                                                             |           |      |        |
| DR | MIM; 602663; -                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |           |      |        |
| DR | GO; GO:0005180; P:peptide hormone; TAS.                                                                                                                                                                                                                                                                                                                                                                                                                                                   |           |      |        |
| DR | Hormone; Amidation; Signal.                                                                                                                                                                                                                                                                                                                                                                                                                                                               |           |      |        |
| KW | SIGNAL 1 22 BY SIMILARITY.                                                                                                                                                                                                                                                                                                                                                                                                                                                                |           |      |        |
| FT | PEPTIDE 23 53 PROLACTIN-RELEASING PEPTIDE PRRP31.                                                                                                                                                                                                                                                                                                                                                                                                                                         |           |      |        |
| FT | PEPTIDE 34 53 PROLACTIN-RELEASING PEPTIDE PRRP20.                                                                                                                                                                                                                                                                                                                                                                                                                                         |           |      |        |
| FT | PROPEP 58 87                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |           |      |        |
| FT | MOD RES 53 53 AMIDATION (G-54 PROVIDE AMIDE GROUP).                                                                                                                                                                                                                                                                                                                                                                                                                                       |           |      |        |
| FT | SEQUENCE 87 AA; 9639 MW; 229A2F3F50CF981B CRC64;                                                                                                                                                                                                                                                                                                                                                                                                                                          |           |      |        |

Q9p5n0 schizosacch  
P18897 rattus norv  
P77562 escherichia  
P54974 agrobacteri  
P30143 escherichia  
P31795 radiation m  
P17812 homo sapien  
P32770 saccharomyc  
P03357 akr murine  
Q9uia9 homo sapien  
Q9epk7 mus musculu  
P03356 akr murine

34 44.5 26.0 1465 1 YH85\_SCHPO  
35 44 25.7 137 1 SNR2\_RAT  
36 44 25.7 364 1 YAIW\_ECOLI  
37 44 25.7 386 1 CRTY\_AGRAU  
38 44 25.7 476 1 YAAJ\_ECOLI  
39 44 25.7 581 1 POL\_MLVRK  
40 44 25.7 591 1 PYRG\_HUMAN  
41 44 25.7 719 1 NRP1\_YEAST  
42 44 25.7 843 1 POL\_MLVAK  
43 44 25.7 1087 1 XPO7\_HUMAN  
44 44 25.7 1087 1 XPO7\_MOUSE  
45 44 25.7 1196 1 POL\_MLVAV

Query Match 100.0%; Score 171; DB 1; Length 87;  
 Best Local Similarity 100.0%; Pred. No. 4.8e-19;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRTHRSMEIRTPDINPAWYASRGIRPVGRF 31  
 |||||  
 DB 23 SRTHRSMEIRTPDINPAWYASRGIRPVGRF 53

## RESULT 2

PRRP BOVIN STANDARD; PRT; 98 AA.  
 AC P81264;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Prolactin-releasing peptide precursor (PrRP) (Prolactin-releasing  
 hormone) [Contains: Prolactin-releasing peptide PrRP31; Prolactin-  
 releasing peptide PrRP20].  
 GN PRH.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OC NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-52.  
 RC TISSUE=Brain;  
 RX MEDLINE=98268781; PubMed=9607765;  
 RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,  
 RA Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,  
 RA Kurokawa T., Nishimura O., Onda H., Fujino M.,  
 RT "A prolactin-releasing peptide in the brain."  
 RL Nature 393:272-276(1998).  
 CC -1- FUNCTION: Stimulates prolactin (PRL) release and regulates the  
 expression of prolactin through its receptor GPR10. May stimulate  
 lactotrophs directly to secrete PRL.  
 CC -1- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AB015417; BAA29025.1; --  
 KW Hormone; Amidation; Signal; Cleavage on pair of basic residues.  
 FT SIGNAL 1 22  
 FT PPTIDE 23 53 Prolactin-releasing peptide PrRP31.  
 FT PPTIDE 33 53 Prolactin-releasing peptide PrRP20.  
 FT PROPEP 58 98  
 FT MOD RES 53 53 AMIDATION (G-54 PROVIDE AMIDE GROUP).  
 SQ SEQUENCE 98 AA; 10544 MW; 08AC35A13B0FA908 CRC64;

Query Match 92.4%; Score 158; DB 1; Length 98;  
 Best Local Similarity 90.3%; Pred. No. 5.1e-17;  
 Matches 28; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SRTHRSMEIRTPDINPAWYASRGIRPVGRF 31  
 |||||  
 DB 23 SRAHQHSMETIRTPDINPAWYASRGIRPVGRF 53

## RESULT 3

PRRP RAT STANDARD; PRT; 83 AA.  
 AC P81278; Q8K3Y0;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Prolactin-releasing peptide precursor (PrRP) (Prolactin-releasing  
 hormone) [Contains: Prolactin-releasing peptide PrRP31; Prolactin-  
 releasing peptide PrRP20].  
 GN PRH.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OC NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Brain;  
 RX MEDLINE=98268781; PubMed=9607765;  
 RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,  
 RA Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,  
 RA Kurokawa T., Nishimura O., Onda H., Fujino M.,  
 RT "A prolactin-releasing peptide in the brain."  
 RL Nature 393:272-276(1998).  
 CC [2]  
 RN SEQUENCE FROM N.A. (ISOFORM 2).  
 RP STRAIN=Sprague-Dawley; TISSUE=Hypothalamus;  
 RC Anderson S.T., Kokay I.C., Lang T., Grattan D.R., Curlewis J.D.;  
 RT "Quantitation of prolactin-releasing peptide (PrRP) mRNA expression in  
 specific brain regions during the rat oestrous cycle and in  
 lactation."  
 RT Lactation."  
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP TISSUE SPECIFICITY.  
 RX MEDLINE=99426652; PubMed=10498338;  
 RA Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Habata Y., Hinuma S.,  
 RA Sekiguchi M., Kitada C., Kurokawa T., Nishimura O., Onda H.,  
 RA Sumino Y., Fujino M.,  
 RT Tissue distribution of prolactin-releasing peptide (PrRP) and its  
 receptor."  
 RL Regul. Pept. 83:1-10(1999).  
 CC -1- FUNCTION: Stimulates prolactin (PRL) release and regulates the  
 expression of prolactin through its receptor GPR10. May stimulate  
 lactotrophs directly to secrete PRL.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event-Alternative splicing; Named isoforms=2;  
 CC Name=1;  
 CC IsoId=P81278-1; Sequences=Displayed;  
 CC Name=2;  
 CC IsoId=P81278-2; Sequences=VSP 004370;  
 CC -1- TISSUE SPECIFICITY: Widely expressed, with highest levels in  
 medulla oblongata and hypothalamus.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AB015418; BAA29026.1; --  
 DR EMBL; AF521930; AAM82154.1; --  
 DR PIR; JC7607; JC7607.  
 KW Hormone; Amidation; Signal; Cleavage on pair of basic residues;  
 KW Alternative splicing.  
 FT SIGNAL 1 21 BY SIMILARITY.  
 FT PPTIDE 22 52 Prolactin-releasing peptide PrRP31.  
 FT PPTIDE 33 52 Prolactin-releasing peptide PrRP20.  
 FT PROPEP 57 83  
 FT MOD RES 52 52  
 FT VARSPLIC 33 83  
 FT AMIDATION (G-53 PROVIDE AMIDE GROUP).  
 FT TPDINPAWYTGIRPVGRFRRATPRDVTGLGQSLCLPL  
 FT DGRTRFSORG -> SECLTYGKPLTSPHPTFSQMP (in  
 isoform 2).  
 FT /FTId=VSP 004370.  
 FT DOC75A264EE84F29 CRC64;  
 SQ SEQUENCE 83 AA; 9215 MW; 9215 MW;  
 Query Match 87.1%; Score 149; DB 1; Length 83;  
 Best Local Similarity 83.9%; Pred. No. 9.8e-16;  
 Matches 26; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

**Query Match** 31.6%; Score 54; DB 1; Length 428;  
**Best Local Similarity** 37.0%; Pred. No. 1.5; Mismatches 6; Indels 0; Gaps 0;

**Matches** 10; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

**Query** 2 RTHRSMSEIRTPDINPAWYASRGIRPV 28  
Db 195 KTRPHSLMIYSDDLGVTHHGLLRPM 221

**RESULT 5**

Y587\_PASMU STANDARD; PRT; 288 AA.

ID Y587\_PASMU  
AC Q9CN56;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein PM0587.  
GN PM0587.  
OS Pasteurella multocida.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
OC Pasteurellaceae; Pasteurella.  
OX NCBI\_TaxID=747;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Pm70;  
RX MEDLINE=21145866; PubMed=11248100;  
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;  
RT "Complete genomic sequence of Pasteurella multocida Pm70.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).  
CC -1- SIMILARITY: BELONGS TO THE FRUCTOSAMINE KINASE FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement/ or send an email to license@isb-sib.ch).

**Query Match** 30.4%; Score 52; DB 1; Length 288;  
**Best Local Similarity** 37.5%; Pred. No. 2; Mismatches 9; Indels 0; Gaps 0;

**Matches** 9; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

**Query** 5 RHRSMSEIRTPDINPAWYASRGIRPV 28  
Db 20 KHKEKIHGEMHEAWIIDGGIOPV 43

**RESULT 6**

ID Y468\_MYCPN STANDARD; PRT; 1882 AA.

AC P75109; Q50317;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Protein MG468 homolog (K05orf1892).

**Query Match** 30.4%; Score 52; DB 1; Length 288;  
**Best Local Similarity** 37.5%; Pred. No. 2; Mismatches 9; Indels 0; Gaps 0;

**Matches** 9; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

**Query** 5 RHRSMSEIRTPDINPAWYASRGIRPV 28  
Db 20 KHKEKIHGEMHEAWIIDGGIOPV 43

**RESULT 6**

ID Y468\_MYCPN STANDARD; PRT; 1882 AA.

AC P75109; Q50317;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Protein MG468 homolog (K05orf1892).





RA Rabbinoiwtsch E., Rajandream M.A., Rutherford K., Rutter S.,  
 RA Seeger K., Saunders D., Sharp S., Squares S., Taylor K.,  
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,  
 RA Hopwood D.A.;  
 RT "Complete genome sequence of the model actinomycete Streptomyces  
 coelicolor A3(2).";  
 RL Nature 417:141-147(2002).  
 CC -!- FUNCTION: BIDIRECTIONALLY DEGRADES SINGLE-STRANDED DNA INTO LARGE  
 CC ACID-INSOLUBLE OLIGONUCLEOTIDES, WHICH ARE THEN DEGRADED FURTHER  
 CC INTO SMALL ACID-SOLUBLE OLIGONUCLEOTIDES (BY SIMILARITY).  
 CC -!- CATALYTIC ACTIVITY: Exonucleolytic cleavage in either 5'- to 3'-  
 CC or 3'- to 5'-direction to yield nucleoside 5'-phosphates.  
 CC -!- SUBUNIT: Heterooligomer composed of large and small subunits (By  
 CC similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -!- SIMILARITY: BELONGS TO THE XSEA FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AL939122; CAC05901.1; -;  
 CC HAMAP; MF\_00378; -; 1.  
 CC InterPro; IPR003753; Exonuc VII\_L.  
 CC DR InterPro; IPR004365; tRNA anti\_L.  
 CC DR Pfam; PF02601; Exonuc VII\_L; 1.  
 CC DR Pfam; PF01336; tRNA anti\_L; 1.  
 CC DR TIGRFAMs; TIGR00237; xsea; 1.  
 CC KW Hydrolyase; Nuclease; Exonuclease; Complete proteome.  
 CC SQ SEQUENCE 402 AA; 43882 MW; 145929A8372B4E08 CRC64;  
 CC  
 CC Query Match 27.5%; Score 47; DB 1; Length 402;  
 CC Best Local Similarity 47.6%; Pred. No. 16;  
 CC Matches 10; Conservative 1; Mismatches 2; Indels 8; Gaps 1;  
 CC  
 CC QY 17 PAMVYASRG-----IRPVG 29  
 CC |||||  
 CC Db 89 PEWYARGQLSRAEIKPVG 109  
 CC |||||  
 CC  
 CC RESULT 11  
 CC EXL1 HUMAN  
 CC ID\_EXL1\_HUMAN STANDARD; PRT; 676 AA.  
 CC AC Q92935;  
 CC DT 15-JUL-1998 (Rel. 36, Created)  
 CC DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 CC DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 CC DE Exostosin-like 1 (EC 2.4.1.224) (Glucuronosyl-N-acetylglucosaminyl-  
 CC DE proteoglycan 4-alpha-N-acetylglucosaminyltransferase) (Exostosin-L)  
 CC DE (Multiple exostosin-like protein).  
 CC GN EXTL1 OR EXTL.  
 CC OS Homo sapiens (Human).  
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 CC OX NCBI\_TaxID=9606;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RX MEDLINE=97189339; PubMed=9037597;  
 CC RA Wise C.A., Clines G.A., Massa H., Trask B.J., Lovett M.;  
 CC RT "Identification and localization of the gene for EXTL1, a third member  
 CC of the multiple exostosins gene family.";  
 CC RL Genome Res. 7:10-16(1997).  
 CC [2]  
 CC RN SEQUENCE FROM N.A., AND VARIANT HIS-379.  
 CC RX MEDLINE=99408231; PubMed=10480354;  
 CC RA Xu L., Xia J., Jiang H., Zhou J., Li H., Wang D., Pan Q., Long Z.,  
 CC RA Fan C., Deng H.-X.;  
 CC RT "Mutation analysis of hereditary multiple exostoses in the Chinese.";  
 CC RL Hum. Genet. 105:45-50(1999).

[3]  
 RN SEQUENCE FROM N.A.  
 RP Wuyts W., Spieker N., Van Roy N., De Paape A., De Boulle K.,  
 RA Willems P.J., Van Hul W., Versteeg R., Speleman F.;  
 RT "Refined physical mapping and genomic structure of the EXTL1 gene.";  
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Probable glycosyltransferase (By similarity).  
 CC -!- CATALYTIC ACTIVITY: UDP-N-acetyl-D-glucosamine + beta-D-  
 CC glucuronosyl-(1->4)-N-acetyl-alpha-D-glucosaminyl-proteoglycan =  
 CC UDP + N-acetyl-alpha-D-glucosaminyl-(1->4)-beta-D-glucuronosyl-  
 CC (1->4)-N-acetyl-alpha-D-glucosaminyl-proteoglycan.  
 CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Endoplasmic  
 CC reticulum (By similarity).  
 CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 47.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; U67191; AAC51141.1; -;  
 CC EMBL; AF083633; AAD02840.1; JOINED.  
 CC EMBL; AF083623; AAD02840.1; JOINED.  
 CC EMBL; AF083624; AAD02840.1; JOINED.  
 CC EMBL; AF083625; AAD02840.1; JOINED.  
 CC EMBL; AF083626; AAD02840.1; JOINED.  
 CC EMBL; AF083627; AAD02840.1; JOINED.  
 CC EMBL; AF083628; AAD02840.1; JOINED.  
 CC EMBL; AF083629; AAD02840.1; JOINED.  
 CC EMBL; AF083630; AAD02840.1; JOINED.  
 CC EMBL; AF083631; AAD02840.1; JOINED.  
 CC EMBL; AF083632; AAD02840.1; JOINED.  
 CC EMBL; AF153980; AAF73172.1; -;  
 CC EMBL; AF153991; AAF73172.1; JOINED.  
 CC Genew; HGNC:3515; EXTL1.  
 CC MIM; 601738; -;  
 CC GO; GO:0008181; P: tumor suppressor; TAS.  
 CC GO; GO:0001501; P: skeletal development; TAS.  
 CC InterPro; IPR004263; Exostosin.  
 CC Pfam; PF03016; Exostosin; 1.  
 CC KW Transferase; Glycosyltransferase; Endoplasmic reticulum;  
 CC KW Transmembrane; Signal-anchor; CYTOPLASMIC (POTENTIAL).  
 CC FT DOMAIN 1 9 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
 CC FT TRANSMEM 10 30 (POTENTIAL).  
 CC FT DOMAIN 31 676 LUMENAL (POTENTIAL).  
 CC FT CARBOHYD 269 269 N-LINKED (GLCNAC... ) (POTENTIAL).  
 CC FT VARIANT 379 379 N -> H.  
 CC FT SEQUENCE 676 AA; 74673 MW; B5E006A8762E5633 CRC64;  
 CC  
 CC Query Match 27.5%; Score 47; DB 1; Length 676;  
 CC Best Local Similarity 45.0%; Pred. No. 29;  
 CC Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;  
 CC  
 CC QY 12 TPDINPAMVYASRGIRPVGRF 31  
 CC |||||  
 CC Db 400 SPQDFPFYLLQGSREPEGRF 419  
 CC |||||  
 CC  
 CC RESULT 12  
 CC PLSC\_HELPY  
 CC ID\_PLSC\_HELPY STANDARD; PRT; 240 AA.  
 CC AC Q25903;  
 CC DT 15-DEC-1998 (Rel. 37, Created)  
 CC DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 CC DE 1-acetyl-sn-glycero-3-phosphate acyltransferase (EC 2.3.1.51) (1-AGP  
 CC DE acyltransferase) (1-AGPAT) (Lysophosphatidic acid acyltransferase)  
 CC DE (LPAAT).

```

GN PLSC OR HP1348.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OC NCBI_TaxID=210;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-P., White O., Kexlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey B.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Matthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori.";
RL Nature 388:539-547 (1997).
CC -1- FUNCTION: CONVERTS LYXOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC
CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION.
CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate =
CC CoA + 1,2-diacyl-sn-glycerol 3-phosphate.
CC -1- PATHWAY: De novo phospholipid biosynthesis; second step.
CC -1- SUBCELLULAR LOCATION: Inner membrane-associated (Potential).
CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE
CC ACYLTRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC ENBL; AEO0636; AAD08393.1; -.
CC PIR; D64688; D64688.
CC TIGR; HP1348; -.
CC InterPro; IPR002123; Acyltransferase.
CC InterPro; IPR004552; AGP acyltin.
CC Pfam; PF01553; Acyltransferase; 1.
CC SMART; SM00563; PlcG; 1.
CC TIGRFAMs; TIGR00530; AGP acyltin; 1.
CC Phospholipid biosynthesis; Transferase; Acyltransferase;
CC Inner membrane; Complete proteome.
CC KW SEQUENCE 240 AA; 27745 MW; 22BD5D08B190BDD CRC64;
CC -----
CC Query Match 27.2%; Score 46.5; DB 1; Length 240;
CC Best Local Similarity 47.6%; pred. No. 11;
CC Matches 10; Conservative 2; Mismatches 8; Indels 1; Gaps 1;
CC -----
QY 1 SRTRHSMRIPTDIN-PAWY 20
DB :|||:|||||
DB 197 ARTRLVMSLESTEDFNSPTWY 217

RESULT 13
POLN HEVBU STANDARD; PRT; 1693 AA.
AC P29374;
DT 01-DEC-1992 (Rel. 24, Created)
DT 28-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Non-structural polyprotein [Contains: RNA-directed RNA polymerase
DE (EC 2.7.7.48); Helicase].
OS Hepatitis E virus (strain Burma) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OC NCBI_TaxID=31767;
RN [1]

```



```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D10330; BAA01172.1; -.
CC DR MEROPS; C41.001; -.
CC DR InterPro; IPR002589; Alpp.
CC DR InterPro; IPR001788; RNA_dep_RNAPol2.
CC DR InterPro; IPR007095; RNA_pol_DS_PS.
CC DR InterPro; IPR007094; RNA_pol_PSVir.
CC DR InterPro; IPR002588; V_methyltransf.
CC DR InterPro; IPR000606; Viral_helicase.
CC DR Pfam; PF01661; Aldp; 1.
CC DR Pfam; PF00978; RNA_dep_RNAPol2; 1.
CC DR Pfam; PF01443; Viral_helicase1; 1.
CC DR Pfam; PF01660; Vmethyltransf; 1.
CC DR SMART; SM00506; Alpp; 1.
CC DR Polyprotein; Transferase; RNA-directed RNA polymerase; Helicase;
CC KW ATP-binding; 975 982 ATP (POTENTIAL).
CC FT NP_BIND 975 982
CC SQ SEQUENCE 1693 AA; 185215 MW; AAB4C9140A7E21EA CRC64;
Query Match 27.2%; Score 46.5; DB 1; Length 1693;
Best Local Similarity 37.0%; Pred. No. 96;
Matches 10; Conservative 5; Mismatches 11; Indels 1; Gaps 1;
QY 2 RTRHSMETRTDINPAWY-ASRGIRP 27
DB 904 RNRPGDELYLPDLAARWFEANRTPR 930
RESULT 15
Y468_MYCGE STANDARD; PRT; 1783 AA.
AC Q49460;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MG468.
GN MG468.
OS Mycoplasma genitalium.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=96026346; PubMed=7569993;
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
RA Nguyen D.T., Uterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-F., Dougherty B.A., Bitt K.F., Hu P.-C., Lucier T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT "The minimal gene complement of Mycoplasma genitalium.";
RL Science 270:397-403(1995).
RN [2]
RP REVISIONS.
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
RA Nguyen D.T., Uterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-F., Dougherty B.A., Bitt K.F., Hu P.-C., Lucier T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 879-985 FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=94075230; PubMed=8253680;

```

```

RA Peterson S.N., Hu P.-C., Bitt K.F., Hutchison C.A. III;
RT "A survey of the Mycoplasma genitalium genome by using random
RT sequencing."; J. Bacteriol. 175:7918-7930(1993).
RL J. Bacteriol. 175:7918-7930(1993).
CC -|- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -|- SIMILARITY: SOME, TO MG064.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U39728; AAC72488.1; -.
CC DR EMBL; U01808; AAD12339.1; -.
CC DR TIGR; MG468; -.
CC DR InterPro; IPR003838; DUF214.
CC DR Pfam; PF02687; PteX; 1.
CC KW Hypothetical protein; Transmembrane; Complete proteome.
CC FT TRANSMEM 16 36
CC FT TRANSMEM 917 937 POTENTIAL.
CC FT TRANSMEM 967 987 POTENTIAL.
CC FT TRANSMEM 1010 1030 POTENTIAL.
CC FT TRANSMEM 1084 1104 POTENTIAL.
CC FT TRANSMEM 1680 1680 POTENTIAL.
CC FT TRANSMEM 1709 1729 POTENTIAL.
CC FT TRANSMEM 1730 1750 POTENTIAL.
CC FT TRANSMEM 1752 1772 POTENTIAL.
CC SQ SEQUENCE 1783 AA; 200168 MW; 87BD575AEC2E374B CRC64;
Query Match 27.2%; Score 46.5; DB 1; Length 1783;
Best Local Similarity 29.0%; Pred. No. 1e+02;
Matches 9; Conservative 6; Mismatches 13; Indels 3; Gaps 1;
QY 1 SRTHSMETRTDINPAWYASRGIRPVGRF 31
DB 1118 RNRPNYKLNQITPESGMYA---IQPYSHF 1145

```

Search completed: December 3, 2003, 19:12:03  
Job time : 7.25 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model  
Run on: December 3, 2003, 19:03:25 ; Search time 26.5 Seconds  
(without alignment)  
301.873 Million cell updates/sec

Title: US-09-868-885b-32  
Perfect score: 171  
Sequence: 1 SRTHSHWEIRTPDINPAWYASRGIRPVGRF 31

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues  
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL 23:.\*  
1: sp archaea:.\*  
2: sp bacteria:.\*  
3: sp fungi:.\*  
4: sp human:.\*  
5: sp invertebrate:.\*  
6: sp mammal:.\*  
7: sp mhc:.\*  
8: sp organelle:.\*  
9: sp phase:.\*  
10: sp plant:.\*  
11: sp rodent:.\*  
12: sp virus:.\*  
13: sp vertebrate:.\*  
14: sp unclassified:.\*  
15: sp virus:.\*  
16: sp bacterioph:.\*  
17: sp archaea:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

| SUMMARIES  |       |             |               | Description        |  |
|------------|-------|-------------|---------------|--------------------|--|
| Result No. | Score | Query Match | Length DB ID  |                    |  |
| 1          | 158   | 92.4        | 98 6 Q8WN12   | Q8WN12 ovis aries  |  |
| 2          | 98    | 57.3        | 117 13 Q9W624 | Q9W624 carassius a |  |
| 3          | 56.5  | 33.0        | 664 16 Q9I1W4 | Q9I1W4 pseudomonas |  |
| 4          | 55    | 32.2        | 315 2 O87474  | O87474 burkholderi |  |
| 5          | 54    | 31.6        | 790 10 Q9M371 | Q9M371 arabidopsis |  |
| 6          | 53.5  | 31.3        | 314 16 O06348 | O06348 mycobacteri |  |
| 7          | 53    | 31.0        | 327 10 Q94K09 | Q94K09 brassica ca |  |
| 8          | 53    | 31.0        | 327 10 Q94F29 | Q94F29 brassica ca |  |
| 9          | 53    | 31.0        | 637 16 Q9ASE9 | Q9ASE9 caulobacter |  |
| 10         | 52    | 30.4        | 785 17 O26276 | O26276 methanobact |  |
| 11         | 51.5  | 30.1        | 420 2 Q33L27  | Q33L27 streptomyce |  |
| 12         | 51    | 29.8        | 54 4 Q9UJF9   | Q9UJF9 homo sapien |  |
| 13         | 51    | 29.8        | 450 11 Q9DA19 | Q9DA19 mus musculu |  |
| 14         | 51    | 29.8        | 465 4 O60687  | O60687 homo sapien |  |
| 15         | 51    | 29.8        | 465 4 Q8W85   | Q8W85 homo sapien  |  |
| 16         | 50    | 29.2        | 123 2 Q45883  | Q45883 eubacterium |  |

|    |      |      |      |    |        |
|----|------|------|------|----|--------|
| 17 | 50   | 29.2 | 128  | 16 | P74747 |
| 18 | 50   | 29.2 | 692  | 2  | Q9L8J6 |
| 19 | 50   | 29.2 | 939  | 5  | Q9GRA0 |
| 20 | 50   | 29.2 | 1253 | 2  | Q8RTV7 |
| 21 | 50   | 29.2 | 1332 | 3  | Q8NXX4 |
| 22 | 50   | 29.2 | 1654 | 16 | Q8FM04 |
| 23 | 49.5 | 28.9 | 176  | 3  | Q08689 |
| 24 | 49   | 28.7 | 72   | 16 | Q9RYB3 |
| 25 | 49   | 28.7 | 173  | 17 | Q97A00 |
| 26 | 49   | 28.7 | 326  | 2  | Q9F642 |
| 27 | 49   | 28.7 | 663  | 16 | Q8PR15 |
| 28 | 49   | 28.7 | 1245 | 2  | Q8KZ57 |
| 29 | 48.5 | 28.4 | 303  | 16 | Q8YWC7 |
| 30 | 48.5 | 28.4 | 436  | 16 | Q8FM14 |
| 31 | 48.5 | 28.4 | 1501 | 10 | Q9SD86 |
| 32 | 48   | 28.1 | 118  | 16 | Q8UKM3 |
| 33 | 48   | 28.1 | 184  | 16 | Q8PER0 |
| 34 | 48   | 28.1 | 220  | 16 | Q91022 |
| 35 | 48   | 28.1 | 335  | 2  | Q45375 |
| 36 | 48   | 28.1 | 335  | 2  | O88002 |
| 37 | 48   | 28.1 | 348  | 5  | O45431 |
| 38 | 48   | 28.1 | 369  | 16 | O8G4X2 |
| 39 | 48   | 28.1 | 455  | 16 | O33340 |
| 40 | 48   | 28.1 | 514  | 5  | Q9VRV3 |
| 41 | 48   | 28.1 | 596  | 5  | Q9VP62 |
| 42 | 48   | 28.1 | 1084 | 5  | Q9TYW4 |
| 43 | 47.5 | 27.8 | 191  | 10 | Q8LMP7 |
| 44 | 47.5 | 27.8 | 252  | 10 | Q8GSK7 |
| 45 | 47.5 | 27.8 | 258  | 11 | Q8BJ19 |

# ALIGNMENTS

## RESULT 1

Q8WN12 ID Q8WN12 PRELIMINARY; PRT; 98 AA.  
AC Q8WN12;  
DT 01-MAR-2002 (TRENBLrel. 20, Created)  
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)  
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)  
DS Preprolactin-releasing peptide.  
OS Ovis aries (Sheep).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;  
OC Bovidae; Caprinae; Ovis.  
OX NCBI\_TaxID=9940;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Curlewis J.D., Kusters D.H.L., Barclay J.L., Anderson S.T.;  
RT "Prolactin-releasing peptide (PrRP) in the ewe: cDNA cloning, mRNA  
RT distribution and effects on prolactin secretion in vitro and in  
RT vivo.";  
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF450453; AAL47178.1;  
SQ SEQUENCE 98 AA; 10513 MW; 2A53331ED62CAAB5 CRC64;

Query Match 92.4%; Score 158; DB 6; Length 98;  
Best Local Similarity 90.3%; Pred. No. 1.4e-15;  
Matches 28; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SRTHSHWEIRTPDINPAWYASRGIRPVGRF 31  
||:|||||||||||||||||||||||  
Db 23 SRTHSHWEIRTPDINPAWYASRGIRPVGRF 53

## RESULT 2

Q9W624 ID Q9W624 PRELIMINARY; PRT; 117 AA.  
AC Q9W624;  
DT 01-NOV-1999 (TRENBLrel. 12, Created)  
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)  
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)

```

DE C-RP amide.
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Carassius.
OX NCBI_TaxID=7957;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Satake H., Minakata H., Fujimoto M.;
RT "Carassius RFamide (C-RF amide).";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB020024; BAA76662.1; -.
SQ SEQUENCE 117 AA; 12879 MW; D5DC4CB22038C2B0 CRC64;

Query Match 57.3%; Score 98; DB 13; Length 117;
Best Local Similarity 53.8%; Pred. No. 1.1e-06;
Matches 14; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

QY 6 HSMETPTDINPAWYASRGIRPVGRF 31
Db 50 HNVNDRSPEDPPWYVGRVPTGRF 75

RESULT 3
QY11W4 PRELIMINARY; PRT; 664 AA.
AC Q911W4;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein PA2151.
GN PA2151.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.B., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964 (2000).
DR EMBL: AB004642; AAG05539.1; -.
DR InterPro; IPR001589; Actbind actnin.
DR Pfam; PF00128; alpha-amylase; 1-.
DR PROSITE; PS00019; ACTININ 1; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 664 AA; 76329 MW; 8F59FEED54C308AD CRC64;

Query Match 33.0%; Score 56.5; DB 16; Length 664;
Best Local Similarity 45.8%; Pred. No. 9.7;
Matches 11; Conservative 4; Mismatches 8; Indels 1; Gaps 1;

QY 4 HRHMEIRTPDINPAWYASRGIRP 27
Db 470 YRPNFFVTDPINP-WFLQSRGP 492

RESULT 4
QY7474 PRELIMINARY; PRT; 315 AA.
ID O87474;
AC O87474;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)

```

```

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Trihydroxytoluene oxygenase.
GN DNTD.
OS Burkholderia cepacia (Pseudomonas cepacia).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DNT;
RX MEDLINE=931194809; PubMed=8449889;
RA Suen W.C., Spain J.C.;
RT "Cloning and characterization of Pseudomonas sp. strain DNT genes for
RT 2,4-dinitrotoluene degradation.";
RL J. Bacteriol. 175:1831-1837 (1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=DNT;
RX MEDLINE=99121037;
RA Haisler B.B., Johnson G.R., Suen W.C., Spain J.C.;
RT "Biochemical and genetic evidence for meta-ring cleavage of 2,4,5-
RT trihydroxytoluene in Burkholderia sp. strain DNT.";
RL J. Bacteriol. 181:965-972 (1999).
DR EMBL: AF076848; AAD12738.1; -.
SQ SEQUENCE 315 AA; 34745 MW; E99261179022961B CRC64;

Query Match 32.2%; Score 55; DB 2; Length 315;
Best Local Similarity 44.4%; Pred. No. 7.1;
Matches 16; Conservative 3; Mismatches 11; Indels 6; Gaps 2;

QY 1 SETHRHSMEIRTPD---INPAWYASRGIRP---VGR 30
Db 216 SRMHCSFEIQDLDAQFLGNKLASRGWKPQGVGR 251

RESULT 5
QY9M371 PRELIMINARY; PRT; 790 AA.
AC Q9M371;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical 87.4 kDa protein.
GN P15G16.60.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsia.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA De Haan M., Maarse A.C., Grivell L.A., Mewes H.W., Lemcke K.,
RA Mayer K.F.X., Quetier F., Salanoubat M.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL132959; CAB71097.1; -.
KW Hypothetical protein.
SQ SEQUENCE 790 AA; 87376 MW; B222724B75690P30 CRC64;

Query Match 31.6%; Score 54; DB 10; Length 790;
Best Local Similarity 52.6%; Pred. No. 27;
Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 13 PDINPAWYASRGIRPVGRF 31
Db 366 PPHNPRYSGRGLQPHGRW 384

RESULT 6
QY06348

```

```

ID O06348 PRELIMINARY; PRT; 314 AA.
AC O06348;
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical 33.2 kDa protein (Oxidoreductase, short-chain
DE dehydrogenase/reductase family).
GN RV3485C OR MT3589 OR MTCV13E12.38C.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RA MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sultston J.B., Taylor K., Whitehead S., Barrell B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
RA Kolonay J.P., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
DR EMBL; Z95390; CAB08708.1; -.
DR EMBL; AE007162; AAK47948.1; ALT_INIT.
DR HSSP; P29132; 1DPI.
DR TIGR; MT3589; -.
DR TubercuList; RV3485C; -.
DR InterPro; IPR002198; ADH short.
DR Pfam; PF00106; adh short; 1.
DR PROSITE; PS00661; ADH short; 1.
KW Hypothetical protein, Oxidoreductase. Complete proteome.
SQ SEQUENCE 314 AA; 33194 MW; 26144BA917B09274 CRC64;

Query Match 31.3%; Score 53.5; DB 16; Length 314;
Best Local Similarity 27.5%; Pred. No. 12;
Matches 11; Conservative 5; Mismatches 11; Indels 13; Gaps 1;

QY 1 SRTHR-----HSMEIRTPDINPAWASRGIRP 27
DB 187 SNTHRFGAYGVTKSAVDHMMKLADELGLGSVVRVNSIRP 226

RESULT 7
ID Q94KU9 PRELIMINARY; PRT; 327 AA.
AC Q94KU9;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Plastid-lipid associated protein PAPI.
OS Brassica campestris (Field mustard).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosidia II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3711;

Query Match 31.0%; Score 53; DB 10; Length 327;
Best Local Similarity 43.8%; Pred. No. 15;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 4 HRHSMIRTPDINPAW 19
DB 45 HRHDFKVRASDVNDWEW 60

RESULT 9
ID Q9A5E9 PRELIMINARY; PRT; 637 AA.
AC Q9A5E9;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Sensory box histidine kinase/response regulator.
GN CC2501.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=155892;
RN SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,

```

```

RN SEQUENCE FROM N.A.
RP MEDLINE=21249173; PubMed=11351096;
RX Kim H.U., Wu S.S., Rathayake C., Huang A.H.;
RT "Brassica rapa Has Three Genes That Encode Proteins Associated with
RT Different Neutral Lipids in Plastids of Specific Tissues.";
RL Plant Physiol. 126:330-341(2001).
DR EMBL; AF290563; AAK57561.1; -.
DR InterPro; IPR006843; PAP_fibrillin.
DR Pfam; PF04755; PAP_fibrillin; 1.
SQ SEQUENCE 327 AA; 35678 MW; 8C7B87FCD6C02422 CRC64;

Query Match 31.0%; Score 53; DB 10; Length 327;
Best Local Similarity 43.8%; Pred. No. 15;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 4 HRHSMIRTPDINPAW 19
DB 45 HRHDFKVRASDVNDWEW 60

RESULT 8
ID Q94FZ9 PRELIMINARY; PRT; 327 AA.
AC Q94FZ9;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Plastid-lipid associated protein PAPI.
OS Brassica campestris (Field mustard).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosidia II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3711;
RN SEQUENCE FROM N.A.
RP MEDLINE=21249173; PubMed=11351096;
RX Kim H.U., Wu S.S., Rathayake C., Huang A.H.;
RT "Brassica rapa Has Three Genes That Encode Proteins Associated with
RT Different Neutral Lipids in Plastids of Specific Tissues.";
RL Plant Physiol. 126:330-341(2001).
DR EMBL; AF290566; AAK57564.1; -.
DR InterPro; IPR006843; PAP_fibrillin.
DR Pfam; PF04755; PAP_fibrillin; 1.
SQ SEQUENCE 327 AA; 35644 MW; 6116E7F1B6C02C88 CRC64;

Query Match 31.0%; Score 53; DB 10; Length 327;
Best Local Similarity 43.8%; Pred. No. 15;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 4 HRHSMIRTPDINPAW 19
DB 45 HRHDFKVRASDVNDWEW 60

RESULT 9
ID Q9A5E9 PRELIMINARY; PRT; 637 AA.
AC Q9A5E9;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Sensory box histidine kinase/response regulator.
GN CC2501.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=155892;
RN SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,

```

```

RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smith K., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
CC -I- SIMILARITY: THE N-TERMINAL REGION IS SIMILAR TO THAT OF OTHER
CC REGULATORY COMPONENTS OF SENSORY TRANSDUCTION SYSTEMS.
CC -I- SIMILARITY: TO PROKARYOTE SENSORY TRANSDUCTION PROTEINS.
DR EMBL: AE005918; AAK24472.1; -.
DR HSSP: Q56312; 3TMV.
DR TIGR: C2501; -.
DR InterPro: IPR003594; ATPbind_ATPase.
DR InterPro: IPR004358; Bact_sens_pr_C.
DR InterPro: IPR003661; His_kinase.
DR InterPro: IPR005467; His_kinase.
DR InterPro: IPR001610; PAC.
DR InterPro: IPR000700; PAS-associ. C.
DR InterPro: IPR000014; PAS_domain.
DR InterPro: IPR001789; Response_reg.
DR Pfam: PF02518; HATPase_c; 1.
DR Pfam: PF00512; Hsika; 1.
DR Pfam: PF00785; PAC; 1.
DR Pfam: PF00989; PAS; 1.
DR Pfam: PF00072; response_reg; 1.
DR PRINTS: PR00344; BCTRLSENSOR.
DR ProDom: PD000039; Response_reg; 1.
DR SMART: SM00387; HATPase_c; 1.
DR SMART: SM00388; Hsika; 1.
DR SMART: SM00086; PAC; 1.
DR SMART: SM00091; PAS; 1.
DR SMART: SM00448; REC; 1.
DR TIGRfam: TIGR00229; sensory_box; 1.
DR PROSITE: PS50109; HIS_KIN; 1.
DR PROSITE: PS50113; PAC; 1.
DR PROSITE: PS50112; PAS; 1.
DR PROSITE: PS50111; PAS; 3.
DR PROSITE: PS50110; RESPONSE_REGULATORY; 1.
DR KINase: Phosphorylation; Sensory transduction; Transferase;
KW Complete proteome.
SQ SEQUENCE 637 AA; 68511 MW; 0EDBEAF76FFA8611 CRC64;

Query Match 31.0%; Score 53; DB 16; Length 637;
Best Local Similarity 48.0%; Pred. No. 30;
Matches 12; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 4 HRHSEIRTPDINPAWYASGIRPV 28
DB 22 HRDSDLSRSPAINPAIRVILRAV 46

RESULT 10
O26276 PRELIMINARY; PRT; 785 AA.
AC O26276;
DT 01-JAN-1998 (TRENBLrel. 05, Created)
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Sensory transduction histidine kinase.
GN MTH174.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Delta H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwan N., Caruso A., Bush D., Safer H., Fatwell D., Prabhakar S.,

```

```

RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
DR EMBL: AB008085; AAB84680.1; -.
DR InterPro: IPR003594; ATPbind_ATPase.
DR InterPro: IPR000515; BPD_transp.
DR InterPro: IPR005467; His_kinase.
DR InterPro: IPR001610; PAC.
DR InterPro: IPR000700; PAS-associ. C.
DR InterPro: IPR000014; PAS_domain.
DR Pfam: PF02518; HATPase_c; 1.
DR Pfam: PF00785; PAC; 1.
DR Pfam: PF00989; PAS; 3.
DR SMART: SM00387; HATPase_c; 1.
DR SMART: SM00086; PAC; 1.
DR SMART: SM00091; PAS; 3.
DR TIGRfam: TIGR00229; sensory_box; 3.
DR PROSITE: PS00402; BPD_TRANSF_INN_MEMBER; 1.
DR PROSITE: PS50109; HIS_KIN; 1.
DR PROSITE: PS50113; PAC; 1.
DR PROSITE: PS50112; PAS; 3.
KW Complete proteome.
SQ SEQUENCE 785 AA; 87726 MW; ADD502C928307986 CRC64;

Query Match 30.4%; Score 52; DB 17; Length 785;
Best Local Similarity 46.4%; Pred. No. 53;
Matches 13; Conservative 1; Mismatches 12; Indels 2; Gaps 1;

QY 4 HRHSEIRTPDINPAWYASGIRPV 29
DB 412 HRATFRRPGNRYWVEYVDRPIRTDG 439

RESULT 11
Q93LZ7 PRELIMINARY; PRT; 420 AA.
AC Q93LZ7;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Chain length factor-like protein.
GN AUR2B.
OS Streptomyces aureofaciens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1894;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CCM3239;
RX MEDLINE=98037514; PubMed=9371463;
RA Kormanec J., Bistakova J., Novakova R., Homerova D., Rezuchova B.;
RT "Cloning and characterization of a new polyketide gene cluster in
RT Streptomyces aureofaciens CCM3239.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY033994; AAK61719.1; -.
DR HSSP: P39435; 1B3N.
DR InterPro: IPR000794; Ketoacyl-synt.
DR Pfam: PF00109; ketoacyl-synt; 1.
DR Pfam: PF02801; ketoacyl-synt C; 1.
DR SEQUENCE 420 AA; 43011 MW; 3C27E22BE88C2DEA CRC64;

Query Match 30.1%; Score 51.5; DB 2; Length 420;
Best Local Similarity 57.9%; Pred. No. 32;
Matches 11; Conservative 2; Mismatches 1; Indels 5; Gaps 1;

QY 18 ANYAS-----RGIRPVGRF 31
DB 37 AWWAAVLRGESGIRPVGRF 55

RESULT 12
Q90JFF

```

```

ID Q9UJF9 PRELIMINARY; PRT; 54 AA.
AC Q9UJF9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE DJ479J7.3 (Sushi-repeat protein (SRPUL)) (Fragment).
GN DJ479J7.3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lawlor S.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL035608; CAB55682.1; -.
FT NON TER 54
SQ SEQUENCE 54 AA; 6110 MW; E2F3C39F7B961A9F CRC64;

Query Match 29.8%; Score 51; DB 4; Length 54;
Best Local Similarity 50.0%; Pred. No. 4;
Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 12 TPDINPAWYASRGIRP 27
Db 18 TPAVPTWYAGSGYYP 33

RESULT 13
ID Q9DA19 PRELIMINARY; PRT; 450 AA.
AC Q9DA19;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE 1700023B02Rik protein.
GN 1700023B02RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CS7BL/6J; TISSUE=Testis;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Glessi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Mateu Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Stauble F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hall D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kotsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK006260; BAB24488.1; -.
DR MGD; MGI:1914185; 1700023B02Rik.
DR InterPro; IPR000345; Cytochrome bind.
DR PROSITE; PS00190; CYTOCHROME_C; 1.
SQ SEQUENCE 450 AA; 51852 MW; F32F11BE6D6A4EAC CRC64;

Query Match 29.8%; Score 51; DB 11; Length 450;
Best Local Similarity 46.7%; Pred. No. 41;

```

```

Matches 14; Conservative 2; Mismatches 8; Indels 6; Gaps 1;

QY 1 SRTRHSHMEIRTPDINPAWYASRGIRPVGR 30
Db 373 SRSHRSHPEKKGSDRN-----RGIRSRSR 396

RESULT 14
O60687 PRELIMINARY; PRT; 465 AA.
AC O60687;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Sushi-repeat protein (Sushi-repeat containing protein).
GN SRPUL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kurosawa H., Inukai T., Inaba T., Goi K., Chang K.-S., Sinjyo T.,
RA Rakestraw K.M., Naeve C.W., Look T.A.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Huang C.-H., Chen H., Peng J., Chen Y.;
RT "Cloning and characterization of the sushi-repeat containing protein (SRP) as a novel interaction partner of Rh type C glycoprotein (RHC3).";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF060567; AAC15765.1; -.
DR EMBL; AF393649; AA073693.1; -.
DR InterPro; IPR001128; Cytochrome_p450.
DR InterPro; IPR003410; Hyalin.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF02494; HYR; 1.
DR Pfam; PF00084; sushi; 3.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
SQ SEQUENCE 465 AA; 52971 MW; 4D752B187FF3EFB8 CRC64;

Query Match 29.8%; Score 51; DB 4; Length 465;
Best Local Similarity 50.0%; Pred. No. 42;
Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 12 TPDINPAWYASRGIRP 27
Db 18 TPAVPTWYAGSGYYP 33

RESULT 15
Q8W85 PRELIMINARY; PRT; 465 AA.
AC Q8W85;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Sushi-repeat protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUS=Placenta;
RA Straubeberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC020733; AA020733.1; -.
DR InterPro; IPR001128; Cytochrome_p450.
DR InterPro; IPR003410; Hyalin.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF02494; HYR; 1.

```

DR Pfam: PF00084; sushi; 3.  
DR SMART: SM00032; CCP; 3.  
DR PROSITE: PS00086; CYTOCHROME\_P450; 1.  
SQ SEQUENCE 465 AA; 52957 MW; 3D7229487DA1B8BD CRC64;  
Query Match 29.8%; Score 51; DB 4; Length 465;  
Best Local Similarity 50.0%; Pred. No. 42;  
Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;  
QY 12 TPDINPAWVASRGIRP 27  
Db ||:|||||  
18 TPAVPTWVAGSGYP 33

Search completed: December 3, 2003, 19:14:02  
Job time : 27.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: December 3, 2003, 18:57:14 ; Search time 33.75 Seconds  
(without alignments)  
145.793 Million cell updates/sec

Title: US-09-868-885B-44

Perfect score: 154

Sequence: 1 SRXHSMEKXTPDINPAWXXRGIRPVGRF 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

A\_Geneseq\_19Jun03:\*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID       | Description        |
|------------|-------|-------------|--------|-------------|--------------------|
| 1          | 147   | 95.5        | 31     | 18 AAW31391 | Human type G prote |
| 2          | 147   | 95.5        | 31     | 18 AAW31384 | Rat type G protein |
| 3          | 147   | 95.5        | 31     | 20 AAW97233 | Rat type ligand po |
| 4          | 147   | 95.5        | 31     | 20 AAW97235 | Human type ligand  |
| 5          | 147   | 95.5        | 31     | 20 AAW87614 | Rat 19P2 ligand.   |
| 6          | 147   | 95.5        | 31     | 20 AAW87615 | Human 19P2 ligand. |
| 7          | 147   | 95.5        | 31     | 20 AAW95173 | Murine pituitary-d |
| 8          | 147   | 95.5        | 31     | 20 AAW95174 | Murine pituitary-d |
| 9          | 147   | 95.5        | 31     | 21 AAB10355 | Rat oxytocin secre |

|    |     |      |    |             |                    |
|----|-----|------|----|-------------|--------------------|
| 10 | 147 | 95.5 | 31 | 21 AAB10362 | Human oxytocin sec |
| 11 | 147 | 95.5 | 31 | 21 AAV87504 | Rat prolactin-rele |
| 12 | 147 | 95.5 | 31 | 21 AAV49291 | 19P2 ligand peptid |
| 13 | 147 | 95.5 | 31 | 21 AAV49292 | 19P2 ligand peptid |
| 14 | 147 | 95.5 | 31 | 22 AAG62524 | Rat CRH releasing  |
| 15 | 147 | 95.5 | 31 | 22 AAG62531 | Human CRH releasin |
| 16 | 147 | 95.5 | 31 | 22 AAB90991 | Proactin releasin  |
| 17 | 147 | 95.5 | 31 | 22 AAB90993 | Proactin releasin  |
| 18 | 147 | 95.5 | 31 | 22 AAB90995 | Proactin releasin  |
| 19 | 147 | 95.5 | 31 | 23 AAE26400 | Rat PRP-31 peptid  |
| 20 | 147 | 95.5 | 31 | 23 AAE26401 | Human PRP-31 pept  |
| 21 | 147 | 95.5 | 31 | 24 ABU60826 | Peptide production |
| 22 | 147 | 95.5 | 31 | 24 ABU60827 | Peptide production |
| 23 | 147 | 95.5 | 31 | 24 ABU60837 | Peptide production |
| 24 | 147 | 95.5 | 31 | 24 ABU60843 | Peptide production |
| 25 | 147 | 95.5 | 32 | 18 AAW31392 | Human type G prote |
| 26 | 147 | 95.5 | 32 | 18 AAW31385 | Rat type G protein |
| 27 | 147 | 95.5 | 32 | 21 AAB10356 | Rat oxytocin secre |
| 28 | 147 | 95.5 | 32 | 21 AAB10363 | Human oxytocin sec |
| 29 | 147 | 95.5 | 32 | 22 AAG62525 | Rat CRH releasing  |
| 30 | 147 | 95.5 | 32 | 22 AAG62532 | Human CRH releasin |
| 31 | 147 | 95.5 | 32 | 24 ABU60838 | Peptide production |
| 32 | 147 | 95.5 | 32 | 24 ABU60844 | Peptide production |
| 33 | 147 | 95.5 | 33 | 18 AAW31393 | Human type G prote |
| 34 | 147 | 95.5 | 33 | 18 AAW31386 | Rat type G protein |
| 35 | 147 | 95.5 | 33 | 21 AAB10357 | Rat oxytocin secre |
| 36 | 147 | 95.5 | 33 | 21 AAB10364 | Human oxytocin sec |
| 37 | 147 | 95.5 | 33 | 22 AAG62526 | Rat CRH releasing  |
| 38 | 147 | 95.5 | 33 | 22 AAG62533 | Human CRH releasin |
| 39 | 147 | 95.5 | 33 | 24 ABU60839 | Peptide production |
| 40 | 147 | 95.5 | 33 | 24 ABU60845 | Peptide production |
| 41 | 147 | 95.5 | 82 | 20 AAW95172 | Murine pituitary-d |
| 42 | 147 | 95.5 | 83 | 18 AAW31383 | Rat type G protein |
| 43 | 147 | 95.5 | 83 | 20 AAW97225 | Rat type ligand po |
| 44 | 147 | 95.5 | 83 | 21 AAB10354 | Rat oxytocin secre |
| 45 | 147 | 95.5 | 83 | 22 AAG62523 | Rat CRH releasing  |

## ALIGNMENTS

### RESULT 1

AAW31391

ID AAW31391 standard; Peptide; 31 AA.

AC AAW31391;

DT 06-APR-1998 (first entry)

DE Human type G protein-coupled receptor ligand fragment 1.

XX G protein-coupled receptor; ligand binding; pharmaceutical;  
modulator; pituitary; central nervous system; pancreas; prophylactic;  
therapeutic agent.

OS Homo sapiens

XX WQ9724436-A2

FN 10-JUL-1997.

PD 26-DEC-1996; 96WO-JP03821.

PP 18-SEP-1996; 96JP-0246573.

PR 28-DEC-1995; 95JP-0343371.

PR 15-MAR-1996; 96JP-0059419.

PR 12-AUG-1996; 96JP-0211805.

XX (TAKE ) TAKEDA CHEM IND LTD.

PA Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M,

PI Kawamata Y, Kitada C;

XX



```
XX Ligand peptide for G protein-coupled receptor - acts by modulating
PT function in the central nervous system, pancreas and pituitary gland
XX
XX Claim 2) Page 179; 258pp; English.
```

---

```
CC This sequence represents a peptide fragment from a novel rat type
CC ligand polypeptide corresponding to amino acid residues 22 to 52 of the
CC sequence represented in AA03183 and is used in an assay to monitor
CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical
CC compositions containing this ligand may be used as a pituitary function
CC modulator, a central nervous system modulator or a pancreatic function
CC modulator. This ligand could have specific applications as a
CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic
CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,
CC trauma, growth hormone secretory disease, hyper- and polyphagia,
CC hyperlipidaemia, hypercholesterolaemia, hypoglycaemia,
CC hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease,
CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,
CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,
CC acute myocardial infarction, infertility, osteoporosis and/or
CC bone fracture, trauma, atopic dermatitis, scoliosis degeneration,
CC oligosaccharia. Assays can also be developed to screen compounds which are
CC capable of altering the binding activity of the ligand affecting
CC activation of the G protein-coupled receptor protein.
XX
```

---

```
XX Sequence      31 AA;
```

---

```
Query Match          95.5%; Score 147; DB 18; Length 31;
Best Local Similarity   83.9%; Pred. No. 3.1e-17;
Matches    26; Conservative     0; Mismatches    5; Indels     0; Gaps
```

---

```
QY       1 SRXHXSMEXRTPDINPAWYKRGIRPVGRF 31
         ||| | |||| | |||| | |||| | |||| | |||| | |||| |
DB        1 SRAHQHMETRPDPINPAAWTGRTGVGRFPVGRF 31
```

---

```
RESULT 3
AANW97233
ID AANW97233 standard; peptide; 31 AA.
```

---

```
AC AANW97233;
XX
DT 06-MAY-1999 (first entry)
```

---

```
DE Rat type ligand polypeptide fragment.
```

---

```
G protein-coupled receptor; prolactin secretion;
KW menopaual syndrome; euthyroid; hypometabolism; lactation;
KW pituitary adenomatosis; brain tumour; emmenopathy; autoimmune disease;
KW prolactinoma; infertility; impotence; amenorrhea; galactorrhea;
KW acromegaly; Chiari-Prommel syndrome; Argonz-del Castillo syndrome;
KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospemia;
KW contraceptive; placental function; choriocarcinoma; hydatid mole;
KW irruption mole; abortion; unthrifty fetus; abnormal saccharomatabolism;
KW abnormal lipidmetabolism; oxytocia.
```

---

```
Rattus sp.
XX
XX WO9858962-A1.
XX
XX PD 30-DEC-1998.
XX
XX PF 22-JUN-1998; 98WO-JP02765.
XX
XX PR 23-JUN-1997; 97JP-0165437.
XX
XX PA (TAKE ) TAKEDA CHEM IND LTD.
XX
XX FI Fujii R, Hinuma S, Kawamata Y, Matsumoto H;
XX WPI; 1999-105614/09.
DR
```

[illegible]



DR WPI; 1999-009423/01.  
 XX  
 PT New polypeptide ligand for orphan G protein coupled receptors - used  
 PT for treating disorders of central nervous system, pituitary and  
 PT pancreas, and for drug screening  
 XX  
 PS Disclosure; Page 134; 206pp; English.  
 XX  
 CC This represents the matured murine pituitary-derived ligand polypeptide  
 CC sequence. The polypeptide is a ligand for the G-protein coupled orphan  
 CC receptor designated GPR10 (human) or UHR-1 (rat). Cells transformed with  
 CC a vector containing the ligand polypeptide encoding DNA are used to  
 CC produce a recombinant ligand polypeptide. The ligand polypeptide, and its  
 CC fragments, modulate function of the pituitary, central nervous system,  
 CC pancreas and other tissues and can be used to screen for agents that  
 CC modulate binding of the polypeptide to the receptor; to quantify the  
 CC amount of receptor in a sample and to raise antibodies. They may also be  
 CC used therapeutically, e.g. to treat senile dementia; Alzheimer's,  
 CC Parkinson's or Huntington's diseases; Creutzfeld-Jakob disease; poisoning  
 CC by heavy metals or drugs; diabetes; schizophrenia; disorders of growth  
 CC hormone secretion; cancer; rheumatoid arthritis, epilepsy and many  
 CC others, also to improve post-operative nutritional status and as  
 CC vasopressor. Transgenic animals carrying the ligand polypeptide encoding  
 CC DNA or its mutein are used to study the function of the polypeptide-  
 CC expressing genes, as models of disease, for drug screening and as source  
 CC of cell lines. The ligand polypeptide DNA is used as a source of probes  
 CC and primers; to identify related sequences; in receptor-binding assays;  
 CC for production of Ab and antisera; in drug development; for gene therapy  
 CC and to develop transgenic animals.  
 XX  
 SQ Sequence 31 AA;  
 Query Match 95.5%; Score 147; DB 20; Length 31;  
 Best Local Similarity 83.9%; Pred. No. 3.3e-17;  
 Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 SRXXHSMEXKTPDINPAWYXXRGIRPVGRF 31  
 |||||  
 DB 1 SRAHQSMETRTPDINPAWYTGRIQVGRF 31  
 |||||  
 RESULT 8  
 AAW95174  
 ID AAW95174 standard; Protein; 31 AA.  
 AC AAW95174;  
 DT 10-MAR-1999 (first entry)  
 XX  
 DE Murine pituitary-derived ligand polypeptide antigenic epitope.  
 XX  
 KW Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;  
 KW GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas;  
 KW tissue; screen; therapeutic; binding; senile dementia; ligand; murine;  
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;  
 KW Creutzfeld-Jakob disease; poisoning; schizophrenia; growth hormone;  
 KW secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;  
 KW gene therapy; transgenic animal; epitope.  
 OS  
 XX Mus sp.  
 XX  
 PN WO9849295-A1.  
 XX  
 PD 05-NOV-1998.  
 XX  
 PF 27-APR-1998; 98WO-JP01923.  
 XX  
 PR 28-APR-1997; 97JP-0109974.  
 XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PI Fukusumi S, Hinuma S;  
 XX

DR WPI; 1999-009423/01.  
 XX  
 PT New polypeptide ligand for orphan G protein coupled receptors - used  
 PT for treating disorders of central nervous system, pituitary and  
 PT pancreas, and for drug screening  
 XX  
 PS Disclosure; Page 26; 206pp; English.  
 XX  
 CC The invention relates to a murine pituitary-derived ligand polypeptide  
 CC which is a ligand for the G-protein coupled orphan receptor designated  
 CC GPR10 (human) or UHR-1 (rat). Cells transformed with a vector containing  
 CC the ligand polypeptide encoding DNA are used to produce a recombinant  
 CC ligand polypeptide. The ligand polypeptide, and its fragments, modulate  
 CC function of the pituitary, central nervous system, pancreas and other  
 CC tissues and can be used to screen for agents that modulate binding of  
 CC the polypeptide to the receptor; to quantify the amount of receptor in a  
 CC sample and to raise antibodies. They may also be used therapeutically,  
 CC e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's  
 CC diseases; Creutzfeld-Jakob disease; poisoning by heavy metals or drugs;  
 CC diabetes; schizophrenia; disorders of growth hormone secretion; cancer;  
 CC rheumatoid arthritis, epilepsy and many others, also to improve post-  
 CC operative nutritional status and as vasopressor. Transgenic animals  
 CC carrying the ligand polypeptide encoding DNA or its mutein are used to  
 CC study the function of the polypeptide-expressing genes, as models of  
 CC disease, for drug screening and as source of cell lines. The ligand  
 CC polypeptide DNA is used as a source of probes and primers; to identify  
 CC related sequences; in receptor-binding assays; for production of Ab and  
 CC antisera; in drug development; for gene therapy and to develop  
 CC transgenic animals. Sequences AAW95174 to AAW95178 represent antigenic  
 CC epitopes which can be used for the preparation of anti-ligand polypeptide  
 CC antibody.  
 XX  
 SQ Sequence 31 AA;  
 Query Match 95.5%; Score 147; DB 20; Length 31;  
 Best Local Similarity 83.9%; Pred. No. 3.3e-17;  
 Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 SRXXHSMEXKTPDINPAWYXXRGIRPVGRF 31  
 |||||  
 DB 1 SRAHQSMETRTPDINPAWYTGRIQVGRF 31  
 |||||  
 RESULT 9  
 AAB10355  
 ID AAB10355 standard; peptide; 31 AA.  
 AC AAB10355;  
 XX  
 DT 24-NOV-2000 (first entry)  
 XX  
 DE Rat oxytocin secretion promoting peptide SEQ ID NO: 18.  
 XX  
 KW Rat; oxytocin secretion promoter; G protein-coupled receptor protein;  
 KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
 KW caesarean section; artificial fertilization; galactostasis; goat; pig;  
 KW veterinary medicine; milk production.  
 XX  
 OS Rattus sp.  
 XX  
 PN WO200038704-A1.  
 XX  
 PD 06-JUL-2000.  
 XX  
 PF 22-DEC-1999; 99WO-JP07199.  
 XX  
 PR 25-DEC-1998; 98JP-0369585.  
 XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PI Mateumoto H, Kitada C, Hinuma S;  
 XX  
 DR WPI; 2000-452298/39.



```

Query Match      95.5%; Score 147; DB 21; Length 31;
Best Local Similarity 83.9%; Pred. No. 3.3e-17;
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 SRXHHSMEXRTPDINPAWYXXRGIRPVGRF 31
DB 1 SRAHQHSMETRTPDINPAWYTGIRPVGRF 31

RESULT 12
AA49291
ID AAY49291 standard; peptide; 31 AA.
XX
AC AAY49291;
XX
DT 22-FEB-2000 (first entry)
XX
DE 19P2 ligand peptide fragment.
XX
KW Monoclonal antibody; 19P2 ligand; diagnosis; prolactin secretion;
KW pituitary; regulatory mechanism; central nervous system; pancreatic.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 31 /note= "C-terminal amide"
XX
PN WO9960112-A1.
XX
PD 25-NOV-1999.
XX
PF 20-MAY-1999; 99WO-JP02650.
XX
PR 21-MAY-1998; 98JP-0140293.
XX
PA (TAKE ) TAKEDA CHEM IND LTD.
XX
PI Matsumoto H, Kitada C, Hinuma S;
XX
DR WPI; 2000-039381/03.
XX
PN WO9960112-A1.
XX
PD 25-NOV-1999.
XX
PF 20-MAY-1999; 99WO-JP02650.
XX
PR 21-MAY-1998; 98JP-0140293.
XX
PA (TAKE ) TAKEDA CHEM IND LTD.
XX
PI Matsumoto H, Kitada C, Hinuma S;
XX
DR WPI; 2000-039381/03.
XX
PN New monoclonal antibodies, useful in diagnosis, as drugs and in
PT studying diseases related to ligand abnormality -
XX
PS Disclosure; Page 26; 73pp; Japanese.
XX
CC The invention provides a monoclonal antibody which has a specific
CC reaction with the part peptide of the C-terminal of 19P2 ligand or its
CC derivative. The antibodies can be used in diagnosis or to treat or
CC prevent diseases associated with abnormality in the pituitary function
CC regulatory mechanism (e.g. promotion of prolactin secretion), central
CC nervous regulatory mechanism, and pancreatic function regulatory
CC mechanism. The antibody-based immunoassay can also be applied in
CC clarifying the physiological functions of the ligand and its derivative.
CC Sequences AAY49290-302 represent peptide fragments of the 19P2 ligand.
XX
SQ Sequence 31 AA;

Query Match      95.5%; Score 147; DB 21; Length 31;
Best Local Similarity 83.9%; Pred. No. 3.3e-17;
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 SRXHHSMEXRTPDINPAWYXXRGIRPVGRF 31
DB 1 SRAHQHSMETRTPDINPAWYTGIRPVGRF 31

RESULT 13
AA49292
ID AAY49292 standard; peptide; 31 AA.
XX
AC AAY49292;
XX
DT 22-FEB-2000 (first entry)
XX
DE 19P2 ligand peptide fragment.
XX
KW Monoclonal antibody; 19P2 ligand; diagnosis; prolactin secretion;
KW pituitary; regulatory mechanism; central nervous system; pancreatic.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 31 /note= "C-terminal amide"
XX
PN WO9960112-A1.
XX
PD 25-NOV-1999.
XX
PF 20-MAY-1999; 99WO-JP02650.
XX
PR 21-MAY-1998; 98JP-0140293.
XX
PA (TAKE ) TAKEDA CHEM IND LTD.
XX
PI Matsumoto H, Kitada C, Hinuma S;
XX
DR WPI; 2000-039381/03.
XX
PN New monoclonal antibodies, useful in diagnosis, as drugs and in
PT studying diseases related to ligand abnormality -
XX
PS Disclosure; Page 26; 73pp; Japanese.
XX
CC The invention provides a monoclonal antibody which has a specific
CC reaction with the part peptide of the C-terminal of 19P2 ligand or its
CC derivative. The antibodies can be used in diagnosis or to treat or
CC prevent diseases associated with abnormality in the pituitary function
CC regulatory mechanism (e.g. promotion of prolactin secretion), central
CC nervous regulatory mechanism, and pancreatic function regulatory
CC mechanism. The antibody-based immunoassay can also be applied in
CC clarifying the physiological functions of the ligand and its derivative.
CC Sequences AAY49290-302 represent peptide fragments of the 19P2 ligand.
XX
SQ Sequence 31 AA;

Query Match      95.5%; Score 147; DB 21; Length 31;
Best Local Similarity 83.9%; Pred. No. 3.3e-17;
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 SRXHHSMEXRTPDINPAWYXXRGIRPVGRF 31
DB 1 SRAHQHSMETRTPDINPAWYTGIRPVGRF 31

RESULT 14
AAG62524
ID AAG62524 standard; peptide; 31 AA.
XX
AC AAG62524;
XX
DT 24-AUG-2001 (first entry)
XX
DE Rat CRH releasing protein related peptide SEQ ID NO: 18.
XX
KW Rat; corticotrophin releasing hormone; CRH; G protein receptor ligand;
KW analgesic; hyperaldosteronism; hypercortisolaemia; hypoadrenocorticism;
KW Addison's disease; adrenal gland hyperfunction; obesity.
XX
OS Rattus sp.
XX
PN WO200135984-A1.
XX
PD 25-MAY-2001.

```

```

XX 22-FEB-2000 (first entry)
XX 19P2 ligand peptide fragment.
XX Monoclonal antibody; 19P2 ligand; diagnosis; prolactin secretion;
XX pituitary; regulatory mechanism; central nervous system; pancreatic.
XX Rattus sp.
XX Key Location/Qualifiers
XX Modified-site 31 /note= "C-terminal amide"
XX WO9960112-A1.
XX 25-NOV-1999.
XX 20-MAY-1999; 99WO-JP02650.
XX 21-MAY-1998; 98JP-0140293.
XX (TAKE ) TAKEDA CHEM IND LTD.
XX Matsumoto H, Kitada C, Hinuma S;
XX WPI; 2000-039381/03.
XX New monoclonal antibodies, useful in diagnosis, as drugs and in
XX studying diseases related to ligand abnormality -
XX Disclosure; Page 26; 73pp; Japanese.
XX The invention provides a monoclonal antibody which has a specific
XX reaction with the part peptide of the C-terminal of 19P2 ligand or its
XX derivative. The antibodies can be used in diagnosis or to treat or
XX prevent diseases associated with abnormality in the pituitary function
XX regulatory mechanism (e.g. promotion of prolactin secretion), central
XX nervous regulatory mechanism, and pancreatic function regulatory
XX mechanism. The antibody-based immunoassay can also be applied in
XX clarifying the physiological functions of the ligand and its derivative.
XX Sequences AAY49290-302 represent peptide fragments of the 19P2 ligand.
XX
SQ Sequence 31 AA;

Query Match      95.5%; Score 147; DB 21; Length 31;
Best Local Similarity 83.9%; Pred. No. 3.3e-17;
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 SRXHHSMEXRTPDINPAWYXXRGIRPVGRF 31
DB 1 SRAHQHSMETRTPDINPAWYTGIRPVGRF 31

RESULT 14
AAG62524
ID AAG62524 standard; peptide; 31 AA.
XX
AC AAG62524;
XX
DT 24-AUG-2001 (first entry)
XX
DE Rat CRH releasing protein related peptide SEQ ID NO: 18.
XX
KW Rat; corticotrophin releasing hormone; CRH; G protein receptor ligand;
KW analgesic; hyperaldosteronism; hypercortisolaemia; hypoadrenocorticism;
KW Addison's disease; adrenal gland hyperfunction; obesity.
XX
OS Rattus sp.
XX
PN WO200135984-A1.
XX
PD 25-MAY-2001.

```

XX PF 17-NOV-2000; 2000WO-JF08119.

XX PR 18-NOV-1999; 99JP-0327900.

XX PR 26-SEP-2000; 2000JP-0297073.

XX PA (TAKE ) TAKEDA CHEM IND LTD.

XX PI Kitada C, Matsumoto H, Hinuma S;

XX DR WPI; 2001-355552/37.

XX PT Use of G protein receptor ligand or peptide for controlling

XX PT corticotropin releasing hormone secretion -

XX PS Claim 3; Page 69; 90pp; Japanese.

XX CC The present sequence describes a method of controlling the secretion of

XX CC corticotrophin releasing hormone (CRH), involving the use of a G protein

XX CC receptor ligand. This can be used to control the secretion of CRH and is

XX CC useful as an analgesic or for treating, preventing or ameliorating

XX CC diseases associated with CRH secretion such as hyperaldosteronism,

XX CC hypercortisolaemia, secondary or chronic hypoadrenocorticism, Addison's

XX CC disease (including boredom, nausea, pigmentation, hypogonadism, hair

XX CC loss, and hypotension), adrenal gland hypofunction and obesity. The

XX CC present sequence is a peptide used in the exemplification of the

XX CC invention.

XX SQ Sequence 31 AA;

Query Match 95.5%; Score 147; DB 22; Length 31;  
Best Local Similarity 83.9%; Pred. No. 3.3e-17;  
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 SRKHCHSXEKTPDINPAWYXKGIRPVGRP 31  
Db 1 SKTRHSMEIKTPDINPAWYASGIRPVGRP 31

Search completed: December 3, 2003, 19:11:24  
Job time : 33.75 secs

|                       |                 |                   |           |            |
|-----------------------|-----------------|-------------------|-----------|------------|
| Query Match           | 95.5%           | Score 147;        | DB 22;    | Length 31; |
| Best Local Similarity | 83.9%           | Pred. NO. 3.e-17; |           |            |
| Matches 26;           | Conservative 0; | Mismatches 5;     | Indels 0; | Gaps 0;    |

**Qy**

1 SRXHXSMEXRTPDINPAWYXXRGIRPVGRF 31  
|| || | ||| | ||| | ||| | ||| | ||| |

**Dd**

1 SRAHQHSMETRTPDINPAWTGRIIPVGRF 31  
|| || | ||| | ||| | ||| | ||| | ||| |

RESULT 15  
AAG62531  
ID AAG62531 standard; peptide: 31 AA.

|    |                                                            |
|----|------------------------------------------------------------|
| XX | AAG62531;                                                  |
| XX | AC                                                         |
| XX | 24-AUG-2001 (first entry)                                  |
| DT |                                                            |
| XX | Human CRH releasing protein related peptide SEQ ID NO: 32. |
| DE |                                                            |

Human; corticotrophin releasing hormone; CRH; G protein receptor ligand;  
KW analgesic; hyperaldosteronism; hypercortisolaemia; hypoadrenocorticism;  
KW Addison's disease; adrenal gland hyperfunction; obesity.

xx  
os  
Homo sapiens.

XX PN WO200135984-A1.

XX  
PD  
25-MAY-2001XX  
17-NOV-2000:

XX  
DB 10-NV-1000-0070 0307000

PR 26-SEP-2000; 2000JP-0297073.

PA (TAKE ) TAKEDA CHEM IND LTD.

PI Kitada C, Matsumoto H, Hinuma S;

DR WPI; 2001-355552/37.

PT Use of G protein receptor ligand or peptide for controlling  
PT corticotropin releasing hormone secretion -

XX  
XXXXXXX releasing hormone secretion -

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 19:09:05 ; Search time 11.5 Seconds  
(without alignment)  
114.055 Million cell updates/sec

Title: US-09-868-885B-44  
Perfect score: 154  
Sequence: 1 SRXHXSMEXRTPDINPAWYXXRGIRPVGRF 31

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A COMB.pcp.\*  
2: /cgn2\_6/ptodata/1/iaa/5B COMB.pcp.\*  
3: /cgn2\_6/ptodata/1/iaa/6A COMB.pcp.\*  
4: /cgn2\_6/ptodata/1/iaa/6B COMB.pcp.\*  
5: /cgn2\_6/ptodata/1/iaa/PCUTS COMB.pcp.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID             | Description       |
|------------|-------|-------------|--------|-------------------|-------------------|
| 1          | 147   | 95.5        | 31     | US-09-105-678A-8  | Sequence 8, Appli |
| 2          | 147   | 95.5        | 31     | US-09-105-678A-9  | Sequence 9, Appli |
| 3          | 147   | 95.5        | 31     | US-09-105-678A-37 | Sequence 37, Appl |
| 4          | 147   | 95.5        | 31     | US-09-105-678A-43 | Sequence 43, Appl |
| 5          | 147   | 95.5        | 31     | US-09-172-353-4   | Sequence 4, Appli |
| 6          | 147   | 95.5        | 31     | US-08-776-971-47  | Sequence 47, Appl |
| 7          | 147   | 95.5        | 31     | US-08-776-971-61  | Sequence 61, Appl |
| 8          | 147   | 95.5        | 31     | US-09-421-208-8   | Sequence 8, Appli |
| 9          | 147   | 95.5        | 31     | US-09-421-208-9   | Sequence 9, Appli |
| 10         | 147   | 95.5        | 31     | US-09-421-208-37  | Sequence 37, Appl |
| 11         | 147   | 95.5        | 31     | US-09-421-208-43  | Sequence 43, Appl |
| 12         | 147   | 95.5        | 31     | US-09-560-915-14  | Sequence 14, Appl |
| 13         | 147   | 95.5        | 31     | US-09-560-915-15  | Sequence 15, Appl |
| 14         | 147   | 95.5        | 31     | US-09-799-955-4   | Sequence 4, Appli |
| 15         | 147   | 95.5        | 32     | US-09-105-678A-38 | Sequence 38, Appl |
| 16         | 147   | 95.5        | 32     | US-09-105-678A-44 | Sequence 44, Appl |
| 17         | 147   | 95.5        | 32     | US-08-776-971-48  | Sequence 48, Appl |
| 18         | 147   | 95.5        | 32     | US-08-776-971-62  | Sequence 62, Appl |
| 19         | 147   | 95.5        | 32     | US-09-421-208-38  | Sequence 38, Appl |
| 20         | 147   | 95.5        | 32     | US-09-421-208-44  | Sequence 44, Appl |
| 21         | 147   | 95.5        | 33     | US-09-105-678A-39 | Sequence 39, Appl |
| 22         | 147   | 95.5        | 33     | US-09-105-678A-45 | Sequence 45, Appl |
| 23         | 147   | 95.5        | 33     | US-08-776-971-49  | Sequence 49, Appl |
| 24         | 147   | 95.5        | 33     | US-08-776-971-63  | Sequence 63, Appl |
| 25         | 147   | 95.5        | 33     | US-09-421-208-39  | Sequence 39, Appl |
| 26         | 147   | 95.5        | 33     | US-09-421-208-45  | Sequence 45, Appl |
| 27         | 147   | 95.5        | 83     | US-08-776-971-45  | Sequence 45, Appl |

Sequence 124, App  
Sequence 137, Appl  
Sequence 59, Appl  
Sequence 135, Appl  
Sequence 138, App  
Sequence 7, Appli  
Sequence 31, Appli  
Sequence 5, Appli  
Sequence 97, Appli  
Sequence 7, Appli  
Sequence 31, Appli  
Sequence 13, Appli  
Sequence 32, Appli  
Sequence 6, Appli  
Sequence 32, Appli  
Sequence 33, Appli  
Sequence 7, Appli  
Sequence 33, Appli

ALIGNMENTS

RESULT 1  
US-09-105-678A-8  
; Sequence 8, Application US/09105678A  
; Patent No. 6103882  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 31 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-8

Query Match 95.5%; Score 147; DB 3; Length 31;  
Best Local Similarity 83.9%; Pred. No. 1.9e-17;  
Matches 26; Conservative 0; Mismatches 5; Indels 0;  
Qy 1 SRXHXSMEXRTPDINPAWYXXRGIRPVGRF 31





```
;
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-105-678A-43

Query Match          95.5%; Score 147; DB 3; Length 31;
Best Local Similarity 83.9%; Pred. No. 1.9e-17;
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 SRXHXHSEKTPDINPAWYXXRGIRPVGRF 31
DB 1 SRTRHSEIRTPDINPAWYASRGIRPVGRF 31

RESULT 5
US-09-172-353-4
; Sequence 4, Application US/09172353
; Patent No. 6197530
; GENERAL INFORMATION:
; APPLICANT: Stricker-Kongra, Alain
; APPLICANT: Gu, Wei
; TITLE OF INVENTION: GPR10 AS A TARGET FOR IDENTIFYING WEIGHT MODULATING COMPOUNDS
; FILE REFERENCE: 07334/102001
; CURRENT APPLICATION NUMBER: US/09/172,353
; CURRENT FILING DATE: 1998-10-14
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-172-353-4

Query Match          95.5%; Score 147; DB 3; Length 31;
Best Local Similarity 83.9%; Pred. No. 1.9e-17;
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 SRXHXHSEKTPDINPAWYXXRGIRPVGRF 31
DB 1 SRAHQHSEKTPDINPAWYXXRGIRPVGRF 31

RESULT 6
US-08-776-971-47
; Sequence 47, Application US/08776971B
; Patent No. 6228984
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Habata, Yugo
; APPLICANT: Kawamata, Yuji
; APPLICANT: Hosoya, Masaki
; APPLICANT: Fujii, Ryo
; APPLICANT: Fukusumi, Shoji
; APPLICANT: Kitada, Chieko
; TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
; NUMBER OF SEQUENCES: 140
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION NUMBER: US/08/776,971B
; FILING DATE: 06-Feb-1997
; CLASSIFICATION: <Unknown>
```

```
;
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/776,971B
; FILING DATE: 06-Feb-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP96/03821
; FILING DATE: 28-DEC-1996
; APPLICATION NUMBER: JP 7/343371
; FILING DATE: 28-DEC-1995
; APPLICATION NUMBER: JP 8/59419
; FILING DATE: 15-MAR-1996
; APPLICATION NUMBER: JP 8/211805
; FILING DATE: 12-AUG-1996
; APPLICATION NUMBER: JP 8/246573
; FILING DATE: 18-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 47176
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 47:
US-08-776-971-47

Query Match          95.5%; Score 147; DB 3; Length 31;
Best Local Similarity 83.9%; Pred. No. 1.9e-17;
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 SRXHXHSEKTPDINPAWYXXRGIRPVGRF 31
DB 1 SRAHQHSEKTPDINPAWYXXRGIRPVGRF 31

RESULT 7
US-08-776-971-61
; Sequence 61, Application US/08776971B
; Patent No. 6228984
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Habata, Yugo
; APPLICANT: Kawamata, Yuji
; APPLICANT: Hosoya, Masaki
; APPLICANT: Fujii, Ryo
; APPLICANT: Fukusumi, Shoji
; APPLICANT: Kitada, Chieko
; TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
; NUMBER OF SEQUENCES: 140
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION NUMBER: US/08/776,971B
; FILING DATE: 06-Feb-1997
; CLASSIFICATION: <Unknown>
```

| Query Match                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | 95.5% | Score 147                      | DB 3 | Length 31 | Best Local Similarity 83.9% | Pred. No. 1.9e-17 | Mismatches 0 | Indels 5 | Gaps 0 |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|--------------------------------|------|-----------|-----------------------------|-------------------|--------------|----------|--------|
| <p>TELECOMMUNICATION INFORMATION:</p> <p>TELEPHONE: 617-523-3400</p> <p>TELEFAX: 617-523-6440</p> <p>INFORMATION FOR SEQ ID NO: 8:</p> <p>SEQUENCE CHARACTERISTICS:</p> <p>LENGTH: 31 amino acids</p> <p>TYPE: amino acid</p> <p>STRANDEDNESS:</p> <p>TOPOLOGY: linear</p> <p>MOLECULE TYPE: peptide</p> <p>US-09-421-208-8</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |       |                                |      |           |                             |                   |              |          |        |
| QY                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | 1     | SRXKHSMEXRTPDINPAWYXXRGIRPVGRF | 31   |           |                             |                   |              |          |        |
| DB                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | 1     | SRXKHSMEXRTPDINPAWYXXRGIRPVGRF | 31   |           |                             |                   |              |          |        |
| <p>RESULT 9</p> <p>US-09-421-208-9</p> <p>Sequence 9, Application US/09421208</p> <p>Patent No. 6258561</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: Suenaga, Masato</p> <p>APPLICANT: Moriya, Takeo</p> <p>APPLICANT: Tanaka, Yoko</p> <p>APPLICANT: Nishimura, Osamu</p> <p>TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND</p> <p>NUMBER OF SEQUENCES: 52</p> <p>CORRESPONDENCE ADDRESS:</p> <p>ADDRESSEE: DIKE, BRONSTEIN, ROBERTS &amp; CUSHMAN, LLP</p> <p>STREET: 130 Water Street</p> <p>CITY: Boston</p> <p>STATE: MA</p> <p>COUNTRY: USA</p> <p>ZIP: 02109</p> <p>COMPUTER READABLE FORM:</p> <p>MEDIUM TYPE: Floppy disk</p> <p>COMPUTER: IBM PC compatible</p> <p>OPERATING SYSTEM: PC-DOS/MS-DOS</p> <p>SOFTWARE: Patent In Release #1.0, Version #1.30</p> <p>CURRENT APPLICATION DATA:</p> <p>APPLICATION NUMBER: US/09/421,208</p> <p>FILING DATE:</p> <p>PRIOR APPLICATION DATA:</p> <p>APPLICATION NUMBER: US 09/105,678</p> <p>FILING DATE: 26-JUN-1998</p> <p>APPLICATION NUMBER: JP 172118/1997</p> <p>FILING DATE: 27-JUN-1997</p> <p>ATTORNEY/AGENT INFORMATION:</p> <p>NAME: Conlin, David G.</p> <p>REGISTRATION NUMBER: 27,026</p> <p>REFERENCE/DOCKET NUMBER: 48466-342</p> <p>TELECOMMUNICATION INFORMATION:</p> <p>TELEPHONE: 617-523-3400</p> <p>TELEFAX: 617-523-6440</p> <p>INFORMATION FOR SEQ ID NO: 9:</p> <p>SEQUENCE CHARACTERISTICS:</p> <p>LENGTH: 31 amino acids</p> <p>TYPE: amino acid</p> <p>STRANDEDNESS:</p> <p>TOPOLOGY: linear</p> <p>MOLECULE TYPE: peptide</p> <p>US-09-421-208-9</p> |       |                                |      |           |                             |                   |              |          |        |
| QY                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | 1     | SRXKHSMEXRTPDINPAWYXXRGIRPVGRF | 31   |           |                             |                   |              |          |        |
| DB                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | 1     | SRXKHSMEXRTPDINPAWYXXRGIRPVGRF | 31   |           |                             |                   |              |          |        |
| <p>Query Match 95.5% Score 147 DB 3 Length 31</p> <p>Best Local Similarity 83.9% Pred. No. 1.9e-17</p> <p>Mismatches 0; Indels 5; Gaps 0</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |       |                                |      |           |                             |                   |              |          |        |
| <p>PRIOR APPLICATION DATA:</p> <p>APPLICATION NUMBER: PCT/JP96/03821</p> <p>FILING DATE: 28-DEC-1996</p> <p>APPLICATION NUMBER: JP 7/343371</p> <p>FILING DATE: 28-DEC-1995</p> <p>APPLICATION NUMBER: JP 8/59419</p> <p>FILING DATE: 15-MAR-1996</p> <p>APPLICATION NUMBER: JP 8/211805</p> <p>FILING DATE: 12-AUG-1996</p> <p>APPLICATION NUMBER: JP 8/246573</p> <p>FILING DATE: 18-SEP-1996</p> <p>ATTORNEY/AGENT INFORMATION:</p> <p>NAME: Conlin, David G.</p> <p>REGISTRATION NUMBER: 27,026</p> <p>REFERENCE/DOCKET NUMBER: 47176</p> <p>TELECOMMUNICATION INFORMATION:</p> <p>TELEPHONE: 617-523-3400</p> <p>TELEFAX: 617-523-6440</p> <p>INFORMATION FOR SEQ ID NO: 61:</p> <p>SEQUENCE CHARACTERISTICS:</p> <p>LENGTH: 31 amino acids</p> <p>TYPE: amino acid</p> <p>STRANDEDNESS: single</p> <p>TOPOLOGY: linear</p> <p>MOLECULE TYPE: protein</p> <p>FRAGMENT TYPE: internal</p> <p>SEQUENCE DESCRIPTION: SEQ ID NO: 61:</p> <p>US-08-776-971-61</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |       |                                |      |           |                             |                   |              |          |        |
| QY                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | 1     | SRXKHSMEXRTPDINPAWYXXRGIRPVGRF | 31   |           |                             |                   |              |          |        |
| DB                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | 1     | SRXKHSMEXRTPDINPAWYXXRGIRPVGRF | 31   |           |                             |                   |              |          |        |
| <p>RESULT 8</p> <p>US-09-421-208-8</p> <p>Sequence 8, Application US/09421208</p> <p>Patent No. 6258561</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: Suenaga, Masato</p> <p>APPLICANT: Moriya, Takeo</p> <p>APPLICANT: Tanaka, Yoko</p> <p>APPLICANT: Nishimura, Osamu</p> <p>TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND</p> <p>NUMBER OF SEQUENCES: 52</p> <p>CORRESPONDENCE ADDRESS:</p> <p>ADDRESSEE: DIKE, BRONSTEIN, ROBERTS &amp; CUSHMAN, LLP</p> <p>STREET: 130 Water Street</p> <p>CITY: Boston</p> <p>STATE: MA</p> <p>COUNTRY: USA</p> <p>ZIP: 02109</p> <p>COMPUTER READABLE FORM:</p> <p>MEDIUM TYPE: Floppy disk</p> <p>COMPUTER: IBM PC compatible</p> <p>OPERATING SYSTEM: PC-DOS/MS-DOS</p> <p>SOFTWARE: Patent In Release #1.0, Version #1.30</p> <p>CURRENT APPLICATION DATA:</p> <p>APPLICATION NUMBER: US/09/421,208</p> <p>FILING DATE:</p> <p>PRIOR APPLICATION DATA:</p> <p>APPLICATION NUMBER: US 09/105,678</p> <p>FILING DATE: 26-JUN-1998</p> <p>APPLICATION NUMBER: JP 172118/1997</p> <p>FILING DATE: 27-JUN-1997</p> <p>ATTORNEY/AGENT INFORMATION:</p> <p>NAME: Conlin, David G.</p> <p>REGISTRATION NUMBER: 27,026</p> <p>REFERENCE/DOCKET NUMBER: 48466-342</p>                                                                                                                                                                                                                                                                                                                                 |       |                                |      |           |                             |                   |              |          |        |

[illegible]



```

STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 32 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-38

Query Match          95.5%; Score 147; DB 3; Length 32;
Best Local Similarity 83.9%; Pred. No. 2e-17;
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      1 SRXKHSMEKRTPDINPAWYXXRGIRPVGRF 31
      ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      1 SRAHQSHMETRPDINPAWYTCGRIRPVGRF 31

Search completed: December 3, 2003, 19:15:55
Job time : 11.5 secs

```

| Result No. | Score | Query |        |    | DB               | ID                | Description |
|------------|-------|-------|--------|----|------------------|-------------------|-------------|
|            |       | Match | Length | †  |                  |                   |             |
| 1          | 147   | 95.5  | 31     | 9  | US-09-932-161-14 | Sequence 14, Appl |             |
| 2          | 147   | 95.5  | 31     | 9  | US-09-932-161-15 | Sequence 15, Appl |             |
| 3          | 147   | 95.5  | 31     | 12 | US-10-096-777-14 | Sequence 14, Appl |             |
| 4          | 147   | 95.5  | 31     | 12 | US-10-096-777-15 | Sequence 15, Appl |             |
| 5          | 147   | 95.5  | 31     | 14 | US-10-044-592-4  | Sequence 4, Appl1 |             |
| 6          | 147   | 95.5  | 31     | 14 | US-10-044-592-5  | Sequence 5, Appl1 |             |
| 7          | 147   | 95.5  | 70     | 14 | US-10-044-592-90 | Sequence 90, Appl |             |
| 8          | 147   | 95.5  | 82     | 14 | US-10-044-592-1  | Sequence 1, Appl1 |             |
| 9          | 147   | 95.5  | 86     | 14 | US-10-044-592-56 | Sequence 56, Appl |             |
| 10         | 147   | 95.5  | 87     | 14 | US-10-044-592-92 | Sequence 92, Appl |             |
| 11         | 147   | 95.5  | 91     | 14 | US-10-044-592-94 | Sequence 94, Appl |             |
| 12         | 146   | 94.8  | 31     | 9  | US-09-932-161-13 | Sequence 13, Appl |             |
| 13         | 146   | 94.8  | 31     | 12 | US-10-096-777-13 | Sequence 13, Appl |             |
| 14         | 146   | 94.8  | 31     | 14 | US-10-044-592-39 | Sequence 39, Appl |             |
| 15         | 146   | 94.8  | 32     | 14 | US-10-044-592-40 | Sequence 40, Appl |             |

RESULT 2  
US-09-932-161-15  
; Sequence 15, Application US/09932161  
; Patent No. US2002003753A1  
; GENERAL INFORMATION:  
; APPLICANT: Civelli, Olivier  
; APPLICANT: Lin, Steven

```
; TITLE OF INVENTION: Screening and Therapeutic Methods For
; FILE REFERENCE: P-UC 4679
; CURRENT APPLICATION NUMBER: US/09/932,161
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 09/560,915
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-932-161-15

Query Match          95.5%; Score 147; DB 9; Length 31;
Best Local Similarity 83.9%; Pred. No. 6.8e-16;
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 SRXHXSMEXRTPDINPAWYXXRGIRPVGRF 31
DB 1 SRTHHSMEIRTPDINPAWYASRGIRPVGRF 31

RESULT 3
US-10-096-777-14
; Sequence 14, Application US/10096777
; Publication No. US20030171270A1
; GENERAL INFORMATION:
; APPLICANT: Civelli, Olivier
; TITLE OF INVENTION: Therapeutic Compositions and Methods
; TITLE OF INVENTION: Relating To Prolactin Releasing Peptide (PrRP)
; FILE REFERENCE: P-UC 3534
; CURRENT APPLICATION NUMBER: US/10/096,777
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: US/09/560,915
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Rattus
US-10-096-777-14

Query Match          95.5%; Score 147; DB 12; Length 31;
Best Local Similarity 83.9%; Pred. No. 6.8e-16;
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 SRXHXSMEXRTPDINPAWYXXRGIRPVGRF 31
DB 1 SRAHQSMETRTDINPAWYTGIRPVGRF 31

RESULT 4
US-10-096-777-15
; Sequence 15, Application US/10096777
; Publication No. US20030171270A1
; GENERAL INFORMATION:
; APPLICANT: Civelli, Olivier
; TITLE OF INVENTION: Therapeutic Compositions and Methods
; TITLE OF INVENTION: Relating To Prolactin Releasing Peptide (PrRP)
; FILE REFERENCE: P-UC 3534
; CURRENT APPLICATION NUMBER: US/10/096,777
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: US/09/560,915
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 31
```

```
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-096-777-15

Query Match          95.5%; Score 147; DB 12; Length 31;
Best Local Similarity 83.9%; Pred. No. 6.8e-16;
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 SRXHXSMEXRTPDINPAWYXXRGIRPVGRF 31
DB 1 SRTHHSMEIRTPDINPAWYASRGIRPVGRF 31

RESULT 5
US-10-044-592-4
; Sequence 4, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 4
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Murine
US-10-044-592-4

Query Match          95.5%; Score 147; DB 14; Length 31;
Best Local Similarity 83.9%; Pred. No. 6.8e-16;
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 SRXHXSMEXRTPDINPAWYXXRGIRPVGRF 31
DB 1 SRAHQSMETRTDINPAWYTGIRPVGRF 31

RESULT 6
US-10-044-592-5
; Sequence 5, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 5
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(31)
```

```
; OTHER INFORMATION: antigen
US-10-044-592-5
Query Match          95.5%; Score 147; DB 14; Length 31;
Best Local Similarity 83.9%; Pred. No. 6.8e-16;
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 SRXHXSMEXRTPDINPAWYXXRGIRPVGRF 31
Db 1 SRAHQSMETRTDINPAWYTGRIQIRPVGRF 31

RESULT 7
US-10-044-592-90
; Sequence 90, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 90
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Rat
US-10-044-592-90

Query Match          95.5%; Score 147; DB 14; Length 70;
Best Local Similarity 83.9%; Pred. No. 1.6e-15;
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 SRXHXSMEXRTPDINPAWYXXRGIRPVGRF 31
Db 22 SRAHQSMETRTDINPAWYTGRIQIRPVGRF 52

RESULT 8
US-10-044-592-1
; Sequence 1, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 1
; LENGTH: 82
; TYPE: PRT
; ORGANISM: Murine
US-10-044-592-1

Query Match          95.5%; Score 147; DB 14; Length 82;
Best Local Similarity 83.9%; Pred. No. 1.8e-15;
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

; OTHER INFORMATION: antigen
US-10-044-592-5
Query Match          95.5%; Score 147; DB 14; Length 31;
Best Local Similarity 83.9%; Pred. No. 6.8e-16;
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 SRXHXSMEXRTPDINPAWYXXRGIRPVGRF 31
Db 21 SRAHQSMETRTDINPAWYTGRIQIRPVGRF 51

RESULT 9
US-10-044-592-96
; Sequence 96, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 96
; LENGTH: 86
; TYPE: PRT
; ORGANISM: mammalian
; NAME/KEY: misc feature
; LOCATION: (1224)..(1243)
; OTHER INFORMATION: Bracket region depicted in FIG 39.
US-10-044-592-96

Query Match          95.5%; Score 147; DB 14; Length 86;
Best Local Similarity 83.9%; Pred. No. 1.9e-15;
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 SRXHXSMEXRTPDINPAWYXXRGIRPVGRF 31
Db 22 SRAHQSMETRTDINPAWYTGRIQIRPVGRF 52

RESULT 10
US-10-044-592-92
; Sequence 92, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 92
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-044-592-92

Query Match          95.5%; Score 147; DB 14; Length 87;
Best Local Similarity 83.9%; Pred. No. 1.9e-15;
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```



QY 1 SRXHXSMEXRTPDINPANYXXRGIRPVGRF 31  
DB 23 SRTHSHMEIRTPDINPANYASRGIRPVGRF 53

## RESULT 11

US-10-044-592-94  
; Sequence 94, Application US/10044592  
; Publication No. US20020143152A1  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; APPLICANT: Fukusumi, Shoji  
; TITLE OF INVENTION: Polypeptides, their Production and Use  
; FILE REFERENCE: 2463US2P  
; CURRENT APPLICATION NUMBER: US/10/044,592  
; PRIOR FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: US 09/403639  
; PRIOR FILING DATE: 1999-25-10  
; PRIOR APPLICATION NUMBER: PCT/JP98/01923  
; PRIOR FILING DATE: 1998-04-27  
; PRIOR APPLICATION NUMBER: JP 9-109974  
; PRIOR FILING DATE: 1997-04-28  
; NUMBER OF SEQ ID NOS: 96  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 94  
; LENGTH: 91  
; TYPE: PRT  
; ORGANISM: Mammalian  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)..(31)  
; OTHER INFORMATION: primer  
; NAME/KEY: misc feature  
; LOCATION: (925)..(955)  
; OTHER INFORMATION: primer  
; NAME/KEY: misc feature  
; LOCATION: (1)..(955)  
; OTHER INFORMATION: insert fragment of pmGB3 as depicted in FIG 36 and 37.  
US-10-044-592-94

Query Match 95.5%; Score 147; DB 14; Length 91;  
Best Local Similarity 83.9%; Pred. No. 2e-15; Mismatches 5; Indels 0; Gaps 0;  
Matches 26; Conservative 0;

QY 1 SRXHXSMEXRTPDINPANYXXRGIRPVGRF 31  
DB 22 SRAHQHSMETRTDINPANYTGRGIRPVGRF 52

## RESULT 12

US-09-932-161-13  
; Sequence 13, Application US/09932161  
; Patent No. US2002037533A1  
; GENERAL INFORMATION:  
; APPLICANT: Civelli, Olivier  
; APPLICANT: Lin, Steven  
; TITLE OF INVENTION: Screening and Therapeutic Methods For  
; FILE REFERENCE: P-UC 4679  
; CURRENT APPLICATION NUMBER: US/09/932,161  
; PRIOR FILING DATE: 2001-08-17  
; PRIOR APPLICATION NUMBER: US 09/560,915  
; PRIOR FILING DATE: 2000-04-28  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 13  
; LENGTH: 31  
; TYPE: PRT  
; ORGANISM: Bos taurus  
US-09-932-161-13

Query Match 94.8%; Score 146; DB 9; Length 31;

Best Local Similarity 83.9%; Pred. No. 9.7e-16;  
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 1 SRXHXSMEXRTPDINPANYXXRGIRPVGRF 31  
DB 1 SRAHQHSMETRTDINPANYAGRGIRPVGRF 31

## RESULT 13

US-10-096-777-13  
; Sequence 13, Application US/10096777  
; Publication No. US20030171270A1  
; GENERAL INFORMATION:  
; APPLICANT: Civelli, Olivier  
; APPLICANT: Lin, Steven  
; TITLE OF INVENTION: Therapeutic Compositions and Methods  
; FILE REFERENCE: P-UC 3534  
; CURRENT APPLICATION NUMBER: US/10/096,777  
; CURRENT FILING DATE: 2002-03-12  
; PRIOR APPLICATION NUMBER: US/09/560,915  
; PRIOR FILING DATE: 2000-04-28  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 13  
; LENGTH: 31  
; TYPE: PRT  
; ORGANISM: Bos taurus  
US-10-096-777-13

Query Match 94.8%; Score 146; DB 12; Length 31;  
Best Local Similarity 83.9%; Pred. No. 9.7e-16;  
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 SRXHXSMEXRTPDINPANYXXRGIRPVGRF 31  
DB 1 SRAHQHSMETRTDINPANYAGRGIRPVGRF 31

## RESULT 14

US-10-044-592-39  
; Sequence 39, Application US/10044592  
; Publication No. US20020143152A1  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; APPLICANT: Fukusumi, Shoji  
; TITLE OF INVENTION: Polypeptides, their Production and Use  
; FILE REFERENCE: 2463US2P  
; CURRENT APPLICATION NUMBER: US/10/044,592  
; CURRENT FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: US 09/403639  
; PRIOR FILING DATE: 1999-25-10  
; PRIOR APPLICATION NUMBER: PCT/JP98/01923  
; PRIOR FILING DATE: 1998-04-27  
; PRIOR APPLICATION NUMBER: JP 9-109974  
; PRIOR FILING DATE: 1997-04-28  
; NUMBER OF SEQ ID NOS: 96  
; SOFTWARE:  
; SEQ ID NO 39  
; LENGTH: 31  
; TYPE: PRT  
; ORGANISM: Bovine  
US-10-044-592-39

Query Match 94.8%; Score 146; DB 14; Length 31;  
Best Local Similarity 83.9%; Pred. No. 9.7e-16;  
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 SRXHXSMEXRTPDINPANYXXRGIRPVGRF 31  
DB 1 SRAHQHSMETRTDINPANYAGRGIRPVGRF 31



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 19:05:30 ; Search time 10.5 Seconds  
(without alignments)  
283.927 Million cell updates/sec

Title: US-09-868-885B-44  
Perfect score: 154  
Sequence: 1 SRHXHSMEXRTPDINPAWYXXRGIRPVGRF 31

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 76:\*  
1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID    | Description        |
|------------|-------|-------------|--------|----------|--------------------|
| 1          | 147   | 95.5        | 83     | 2 JC7607 | prolactin-releasin |
| 2          | 49.5  | 32.1        | 664    | 2 F83376 | conserved hypothet |
| 3          | 49    | 31.8        | 220    | 2 C83292 | probable glutathio |
| 4          | 49    | 31.8        | 790    | 2 T47959 | hypothetical prote |
| 5          | 47    | 30.5        | 128    | 2 S76955 | hypothetical prote |
| 6          | 46    | 29.9        | 118    | 2 AC3169 | hypothetical prote |
| 7          | 46    | 29.9        | 314    | 2 B70569 | hypothetical prote |
| 8          | 46    | 29.9        | 333    | 2 H82852 | hydroxybenzoate oc |
| 9          | 46    | 29.9        | 348    | 2 T21648 | hypothetical prote |
| 10         | 46    | 29.9        | 798    | 2 S11210 | probable unr prote |
| 11         | 46    | 29.9        | 1236   | 2 T50904 | Mg protoporphyrin  |
| 12         | 46    | 29.9        | 1292   | 2 T31462 | probable magnesium |
| 13         | 45    | 29.2        | 240    | 2 B75318 | ferripyochelin-bin |
| 14         | 45    | 29.2        | 501    | 2 T48336 | hypothetical prote |
| 15         | 45    | 29.2        | 637    | 2 D87559 | sensory box histid |
| 16         | 45    | 29.2        | 767    | 2 T21969 | hypothetical prote |
| 17         | 44.5  | 28.9        | 240    | 2 D46888 | probable 1-acylgly |
| 18         | 44.5  | 28.9        | 954    | 2 E83162 | hypothetical prote |
| 19         | 44    | 28.6        | 72     | 2 A91002 | probable regulator |
| 20         | 44    | 28.6        | 159    | 2 G82669 | ubiquinone biosynt |
| 21         | 44    | 28.6        | 250    | 2 G83400 | hypothetical prote |
| 22         | 44    | 28.6        | 284    | 2 F71015 | hypothetical prote |
| 23         | 44    | 28.6        | 548    | 2 T47548 | hypothetical prote |
| 24         | 44    | 28.6        | 962    | 2 H69157 | exonuclease ABC c  |
| 25         | 43.5  | 28.2        | 503    | 2 A82193 | Sun/nucleolar prot |
| 26         | 43.5  | 28.2        | 852    | 2 T28790 | hypothetical prote |
| 27         | 43.5  | 28.2        | 1501   | 2 T45623 | hypothetical prote |
| 28         | 43    | 27.9        | 232    | 2 G75608 | hypothetical prote |
| 29         | 43    | 27.9        | 309    | 2 T32376 | hypothetical prote |

|    |      |      |      |          |                    |
|----|------|------|------|----------|--------------------|
| 30 | 43   | 27.9 | 342  | 2 F91007 | probable gentisate |
| 31 | 43   | 27.9 | 342  | 2 H85851 | probable 1,2-dioxy |
| 32 | 43   | 27.9 | 376  | 2 A48197 | opsin, ocular - A  |
| 33 | 43   | 27.9 | 376  | 2 E48197 | opsin, lateral eye |
| 34 | 43   | 27.9 | 455  | 2 D70885 | probable aldc prot |
| 35 | 43   | 27.9 | 476  | 2 G64720 | probable amino aci |
| 36 | 43   | 27.9 | 476  | 2 G90629 | probable inner mem |
| 37 | 43   | 27.9 | 476  | 2 G85480 | inner membrane tra |
| 38 | 43   | 27.9 | 581  | 2 A42743 | pol polyprotein -  |
| 39 | 43   | 27.9 | 719  | 2 S61046 | ARPI protein - yea |
| 40 | 43   | 27.9 | 843  | 1 GNMVK  | pol polyprotein -  |
| 41 | 43   | 27.9 | 1196 | 1 GNMVGV | HIV-1 retropepsin  |
| 42 | 43   | 27.9 | 1196 | 1 GNMVRV | HIV-1 retropepsin  |
| 43 | 43   | 27.9 | 1296 | 2 T16859 | hypothetical prote |
| 44 | 42.5 | 27.6 | 303  | 2 AH2016 | hypothetical prote |
| 45 | 42.5 | 27.6 | 443  | 2 T21499 | hypothetical prote |

ALIGNMENTS

RESULT 1

JC7607  
prolactin-releasing peptide - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 30-Jun-2001  
C:Accession: JC7607  
F:Yamada, M.; Ozawa, A.; Ishii, S.; Shibusawa, N.; Hashida, T.; Hosoya, T  
Biochem. Biophys. Res. Commun. 281, 53-56, 2001  
A:Title: Isolation and characterization of the rat prolactin-releasing peptide gene: Mu  
A:Reference number: JC7607; MUID:21092785; PMID:11178959  
A:Contents: Spleen  
A:Accession: JC7607  
A:Molecule type: DNA  
A:Residues: 1-83 <YAM>  
A:Cross-references: DDBJ:AB040612; DDBJ:AB040613  
C:Comment: This peptide induces arachidonic acid metabolite release from rat anterior p  
release, and stimulation of ACTH secretion from the pituitary.  
C:Genetics:  
A:Gene: PrRP  
A:Introns: 33/1

|                       |       |                                |      |              |
|-----------------------|-------|--------------------------------|------|--------------|
| Query Match           | 95.5% | Score 147                      | DB 2 | Length 83    |
| Best Local Similarity | 83.9% | Pred. No. 2.5e-16              |      |              |
| Matches               | 26    | Conservative                   | 0    | Mismatches 5 |
|                       |       |                                |      | Indels 0     |
|                       |       |                                |      | Gaps 0       |
| Qy                    | 1     | SRHXHSMEXRTPDINPAWYXXRGIRPVGRF | 31   |              |
|                       |       |                                |      |              |
| Db                    | 22    | SRHXHSMEXRTPDINPAWYXXRGIRPVGRF | 52   |              |

RESULT 2

F83376  
conserved hypothetical protein PA2151 [imported] - Pseudomonas aeruginosa (strain PA01)  
C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: F83376  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim  
; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path  
A:Reference number: AB2950; MUID:20437337; PMID:10984043  
A:Accession: F83376  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-664 <STO>  
A:Cross-references: GB:AE004642; GB:AE004091; NID:G9948163; PIDN:AG05539.1; GSPDB:GN00  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA2151

Query Match 32.1%; Score 49.5; DB 2; Length 664;

Best Local Similarity 62.5%; Pred. No. 8.1;  
Matches 10; Conservative 1; Mismatches 4; Indels 1; Gaps 1;  
QY 12 TPDINPANYXXRGIRP 27  
|||||:|  
Db 478 TPDINP-WFLQSRGP 492  
|||:|  
RESULT 3  
CS3292  
probable glutathione S-transferase PA2821 [imported] - Pseudomonas aeruginosa (strain PA  
C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 23-Mar-2001  
C:Accession: C83292  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bz  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kae, A.; Larbig, K.; Lim,  
Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
A:Reference number: A82950; MUID:20437337; PMID:10984043  
A:Accession: C83292  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-220 <STO>  
A:Cross-references: GB:AE004709; GB:AE004091; NID:g9948304; PIDN:AAG06209.1; GSPDB:GN001  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA2821  
C:Superfamily: plaiace glutathione transferase  
Query Match 31.8%; Score 49; DB 2; Length 220;  
Best Local Similarity 44.0%; Pred. No. 3.1;  
Matches 11; Conservative 3; Mismatches 9; Indels 2; Gaps 1;  
QY 6 HSMEXRTDPINPAWYXXRGIRPVGR 30  
:|:|:|:|:|:|:|:|:|  
Db 28 YQLEAIAPGQAWY-REISPLGR 50  
:|:|:|:|:|:|:|:|:|  
RESULT 4  
T47959  
hypothetical protein F15G16.60 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
C:Accession: T47959  
R:De Haan, M.; Maarse, A.C.; Grivell, L.A.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Que  
submitted to the Protein Sequence Database, January 2000  
A:Reference number: Z24480  
A:Accession: T47959  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-790 <DEH>  
A:Cross-references: EMBL:AL132959  
A:Experimental source: cultivar Columbia; BAC clone F15G16  
C:Genetics:  
A:Map position: 3  
A:Introns: 39/1; 678/2; 698/3; 773/2  
A:Note: F15G16.60  
Query Match 31.8%; Score 49; DB 2; Length 790;  
Best Local Similarity 47.4%; Pred. No. 12;  
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;  
QY 13 PDINPAWYXXRGIRPVGRF 31  
|:|:|:|:|:|:|:|:|  
Db 366 PPHNPRTYSGRGLQPHGRW 384  
|:|:|:|:|:|:|:|:|  
RESULT 5  
S76955  
hypothetical protein - Synechocystis sp. (strain PCC 6803)  
C:Species: Synechocystis sp.  
A:Variety: PCC 6803  
C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999  
C:Accession: S76955  
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.  
O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasu  
DNA Res. 3, 109-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocysti  
s.  
A:Reference number: S74322; MUID:97061201; PMID:8905231  
A:Accession: S76955  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-128 <KAN>  
A:Cross-references: EMBL:D90917; GB:AB001339; NID:G1653836; PIDN:BAAL8867.1; PID:dl0196  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
Query Match 30.5%; Score 47; DB 2; Length 128;  
Best Local Similarity 33.3%; Pred. No. 3.8;  
Matches 11; Conservative 5; Mismatches 13; Indels 4; Gaps 1;  
QY 1 SRXHXSMEXRTDPINPAWYXXR---GIRPVG 29  
:|:|:|:|:|:|:|:|:|  
Db 11 SPENNQLDCSSPDPSRKWYQARLFSFGITPTG 43  
:|:|:|:|:|:|:|:|:|  
RESULT 6  
AC3169  
hypothetical protein Atus077 [imported] - Agrobacterium tumefaciens (strain C58, Dupont  
C:Species: Agrobacterium tumefaciens  
C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 18-Nov-2002  
C:Accession: AC3169  
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo,  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McCle  
Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
ster, E.W.  
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A:Reference number: AB2577; MUID:21608550; PMID:11743193  
A:Accession: AC3169  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-118 <KUR>  
A:Cross-references: GB:AB008687; PIDN:AAL45769.1; PID:gi7743503; GSPDB:GN00188  
A:Experimental source: strain C58 (Dupont)  
C:Genetics:  
A:Gene: Atus077  
A:Genome: plasmid  
Query Match 29.9%; Score 46; DB 2; Length 118;  
Best Local Similarity 34.4%; Pred. No. 5;  
Matches 11; Conservative 4; Mismatches 13; Indels 4; Gaps 1;  
QY 2 RXHXSMEXRTPDI---NPWYXXRGIRPVG 29  
:|:|:|:|:|:|:|:|:|  
Db 19 KLHKNAARRRPDIGARVQAPATAPHRGLAGIG 50  
:|:|:|:|:|:|:|:|:|  
RESULT 7  
B70569  
hypothetical protein Rv3485c - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
C:Accession: B70569  
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S  
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S  
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A:Reference number: A70500; MUID:98295987; PMID:9634230  
A:Accession: B70569  
A:Status: preliminary  
A:Molecule type: DNA

A;Residues: 1-314 <COL>  
A;Cross-references: GB:Z95390; GB:AL123456; NID:g3261766; PIDN:CAB08708.1; PID:g2104408  
A;Experimental source: strain H37RV  
C;Genetics:  
A;Gene: Rv3485c  
C;Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology  
P;46-227/Domain: short-chain alcohol dehydrogenase homology <SADH>

Query Match 29.9%; Score 46; DB 2; Length 314;  
Best Local Similarity 31.8%; Pred. No. 14;  
Matches 7; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

Qy 6 HSMEXRTPDINPAWYXXRGIRP 27  
Db 205 HMKLADELGFSWVRVNSIRP 226

RESULT 8  
H82852  
hydroxybenzoate octaprenyltransferase XF0068 [imported] - Xylella fastidiosa (strain 9a5)  
C;Species: Xylella fastidiosa  
C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 17-Nov-2000  
C;Accession: H82852  
R;anonymus, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequences  
Nature 406, 151-157, 2000  
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A;Reference number: A82515; MUID:20365717; PMID:10910347  
A;Note: for a complete list of authors see reference number A59328 below  
A;Accession: H82852  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-333 <SIM>  
A;Cross-references: GB:A8003860; GB:AE003849; NID:g9104830; PIDN:AAF82881.1; GSPDB:GN001  
A;Experimental source: strain 9a5c  
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; B  
Rimones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrier, H  
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.  
submitted to GenBank, June 2000  
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm  
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig  
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E  
A;Authors: Martins, E.M.F.; Matukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.  
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A  
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak  
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir  
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z  
A;Contents: annotation  
A;Genetics:  
C;Gene: XF0068  
C;Superfamily: 4-hydroxybenzoate octaprenyltransferase

Query Match 29.9%; Score 46; DB 2; Length 333;  
Best Local Similarity 53.3%; Pred. No. 15;  
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 15 INPAWYXXRGIRPVG 29  
Db 54 LDPYWKLAGDRPVG 68

RESULT 9  
T21648  
hypothetical protein F32B6.3 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C;Accession: T21648  
R;Basham, V.  
submitted to the EMBL Data Library, October 1996  
A;Reference number: Z19453  
A;Accession: T21648  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA

A;Residues: 1-348 <WIL>  
A;Cross-references: EMBL:Z81074; PIDN:CAB03039.1; GSPDB:GN00022; CESP:F32B6.3  
A;Experimental source: clone F32B6  
C;Genetics:  
A;Gene: CESP:F32B6.3  
A;Map position: 4  
A;Introns: 56/3; 105/1; 146/3; 185/1; 223/1; 258/2; 320/3

Query Match 29.9%; Score 46; DB 2; Length 348;  
Best Local Similarity 38.9%; Pred. No. 15;  
Matches 7; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

Qy 2 RXHXHSMEXRTPDINPAW 19  
Db 112 RXRLHQLELAQPDNCGW 129

RESULT 10  
S11210  
Probable untr protein - rat  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 21-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 05-Nov-1999  
C;Accession: S11210  
R;Jeffers, M.; Paciucci, R.; Pellicer, A.  
Nucleic Acids Res. 18, 4891-4899, 1990  
A;Title: Characterization of unr; a gene closely linked to N-ras.  
A;Reference number: S11210; MUID:90370473; PMID:2204029  
A;Accession: S11210  
A;Molecule type: mRNA  
A;Residues: 1-798 <JBP>  
A;Cross-references: EMBL:X5311; NID:g57454; PIDN:CAA36549.1; PID:g57455  
C;Keywords: DNA binding

Query Match 29.9%; Score 46; DB 2; Length 798;  
Best Local Similarity 39.1%; Pred. No. 36;  
Matches 9; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

Qy 6 HSMEXRTPDINPAWYXXRGIRP 28  
Db 583 HSVGITEANPTIYSGKVRPL 605

RESULT 11  
T50904  
Mg protoporphyrin methyl transferase [imported] - Rubrivivax gelatinosus  
C;Species: Rubrivivax gelatinosus  
C;Date: 21-Jul-2000 #sequence\_revision 21-Jul-2000 #text\_change 02-Sep-2000  
C;Accession: T50904  
R;Nagashima, K.V.; Igarashi, N.; Harada, J.; Nagashima, S.; Matsuura, K.; Shimada, K.  
submitted to the EMBL Data Library, November 1999  
A;Description: Determination of Nucleotide Sequences of Rubrivivax gelatinosus Photosyn  
A;Reference number: Z25270  
A;Accession: T50904  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-1236 <NAG>  
A;Cross-references: EMBL:AB034704; PIDN:BAA94057.1  
A;Experimental source: strain IL114  
C;Genetics:  
A;Gene: bchH  
C;Superfamily: Rhodobacter capsulatus magnesium-protoporphyrin O-methyltransferase

Query Match 29.9%; Score 46; DB 2; Length 1236;  
Best Local Similarity 32.1%; Pred. No. 56;  
Matches 9; Conservative 5; Mismatches 8; Indels 6; Gaps 1;

Qy 7 SMEXRTPDINPAWY-----XXRGIRP 28  
Db 1116 ALETRRLNPKWEGMLEHGYGVROI 1143

RESULT 12  
T31462

probable magnesium chelatase (EC 4.99.1.-) chain H BchH - *Hellobacillus mobilis*  
 C:Species: *Hellobacillus mobilis*  
 C:Date: 02-Sep-2000 #sequence\_revision 02-Sep-2000 #text\_change 15-Sep-2000  
 C:Accession: T31462  
 R:Xiong, J.; Inoue, K.; Bauer, C.E.  
 Proc. Natl. Acad. Sci. U.S.A. 95, 14851-14856, 1998  
 A:Title: Tracking molecular evolution of photosynthesis by characterization of a major P  
 A:Reference number: Z21036; MUID:99061957; PMID:9843979  
 A:Accession: T31462  
 A:Status: preliminary; translated from GB/EMBL/DBSJ  
 A:Molecule type: DNA  
 A:Residues: 1-1292 <XIO>  
 A:Cross-references: EMBL:AF080002; NID:g3820536; PID:g3820560; PIDN:AAC84033.1  
 C:Genetics:  
 A:Gene: bchH  
 C:Superfamily: Rhodobacter capsulatus magnesium-protoporphyrin O-methyltransferase  
 C:Keywords: lyase

Query Match 29.9%; Score 46; DB 2; Length 1292;  
 Best Local Similarity 34.6%; Pred. No. 59;  
 Matches 9; Conservative 3; Mismatches 8; Indels 6; Gaps 1;

QY 9 EXRTPDINPAWY-----XXRGIRPV 28  
 Db 1176 ETRTTLNFKWYEGMLKHGVGVREI 1201

RESULT 13  
 B75318  
 ferripyochelin-binding protein - *Deinococcus radiodurans* (strain R1)  
 C:Species: *Deinococcus radiodurans*  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Mar-2000  
 C:Accession: B75318  
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
 M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma  
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
 Science 286, 1571-1577, 1999  
 A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.  
 A:Reference number: A75250; MUID:20036896; PMID:10567266  
 A:Accession: B75318  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-240 <WHI>  
 A:Cross-references: GB:AE002044; GB:AE000513; NID:g6459872; PIDN:AAF11635.1; PID:g645988  
 A:Experimental source: strain R1  
 C:Genetics:  
 A:Gene: DR2089  
 A:Map position: 1

Query Match 29.2%; Score 45; DB 2; Length 240;  
 Best Local Similarity 32.0%; Pred. No. 15;  
 Matches 8; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 6 HSMEXRTPDINPAWYXXRGIRPVGR 30  
 Db 63 HTLEDLTPDIHPTAFVAPSADVICQ 87

RESULT 14  
 T48336  
 hypothetical protein F15A17.180 - *Arabidopsis thaliana*  
 C:Species: *Arabidopsis thaliana* (mouse-ear cress)  
 C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
 C:Accession: T48336  
 R:Bevan, M.; Terry, N.; Ardiles, W.; Buysaert, C.; Dasseville, R.; De Clerck, R.; De  
 ewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.P.X.  
 submitted to the Protein Sequence Database, April 2000  
 A:Reference number: Z24491  
 A:Accession: T48336  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-501 <BEV>  
 A:Cross-references: EMBL:AL163002

A:Experimental source: cultivar Columbia; BAC clone F15A17  
 C:Genetics:  
 A:Map position: 5  
 A:Introns: 63/1; 194/2; 280/1  
 A>Note: F15A17.180

Query Match 29.2%; Score 45; DB 2; Length 501;  
 Best Local Similarity 42.9%; Pred. No. 32;  
 Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 4 HXHSMEKXTPDINP 17  
 Db 24 HLHLHQOQIPDLNP 37

## RESULT 15

B7559  
 sensory box histidine kinase/response regulator [imported] - *Caulobacter crescentus*  
 C:Species: *Caulobacter crescentus*  
 C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
 C:Accession: B7559  
 R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kol  
 n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
 A:Title: Complete Genome Sequence of *Caulobacter crescentus*.  
 A:Reference number: A87249; MUID:21173698; PMID:11259647  
 A:Accession: B7559  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-637 <STO>  
 A:Cross-references: GB:AE005673; NID:gl3424056; PIDN:AAK24472.1; GSPDB:GN00148  
 C:Genetics:  
 A:Gene: CC2501

Query Match 29.2%; Score 45; DB 2; Length 637;  
 Best Local Similarity 44.0%; Pred. No. 41;  
 Matches 11; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY 4 HXHSMEKXTPDINPAWYXXRGIRPV 28  
 Db 22 HRDSDDLRSPAINPAAIRVILRAV 46

Search completed: December 3, 2003, 19:14:57  
 Job time : 11.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 19:01:34 ; Search time 6.25 Seconds  
(without alignment)  
233.252 Million cell updates/sec

Title: US-09-868-885B-44

Perfect score: 154

Sequence: 1 SRXHSMEKXRTDPINPAWYXRGIRPVGRF 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID           | Description        |
|------------|-------|-------------|--------|--------------|--------------------|
| 1          | 147   | 95.5        | 83     | 1 PRRP RAT   | P81278 rattus norv |
| 2          | 147   | 95.5        | 87     | 1 PRRP HUMAN | P81277 homo sapien |
| 3          | 146   | 94.8        | 98     | 1 PRRP BOVIN | P81264 bos taurus  |
| 4          | 46    | 29.9        | 676    | 1 EXL1_HUMAN | Q92935 homo sapien |
| 5          | 46    | 29.9        | 798    | 1 UNR RAT    | P83935 rattus norv |
| 6          | 45    | 29.2        | 428    | 1 NER3_BOVIN | O97859 bos taurus  |
| 7          | 44.5  | 28.9        | 240    | 1 PLSC_HELPY | Q25903 helicobacte |
| 8          | 44    | 28.6        | 288    | 1 Y587_PASMU | Q9cn56 pasteurella |
| 9          | 44    | 28.6        | 428    | 1 NER3_HUMAN | Q9uq49 homo sapien |
| 10         | 44    | 28.6        | 962    | 1 UVRA_METTH | O36543 methanobact |
| 11         | 43    | 27.9        | 376    | 1 OPS1_LIMPO | P35360 limulus pol |
| 12         | 43    | 27.9        | 376    | 1 OPS2_LIMPO | P35361 limulus pol |
| 13         | 43    | 27.9        | 383    | 1 CYCR_CHRVI | O82947 chromatium  |
| 14         | 43    | 27.9        | 476    | 1 YAAJ_ECOLI | P30143 escherichia |
| 15         | 43    | 27.9        | 581    | 1 POL_MLRK   | P31795 radiation m |
| 16         | 43    | 27.9        | 719    | 1 NRPI_YEAST | P32770 saccharomyc |
| 17         | 43    | 27.9        | 843    | 1 POL_MLVAK  | P30356 akr murine  |
| 18         | 43    | 27.9        | 1196   | 1 POL_MLVAV  | P30356 akr murine  |
| 19         | 43    | 27.9        | 1196   | 1 POL_MLVAV  | P11227 radiation m |
| 20         | 42.5  | 27.6        | 772    | 1 LMBT_HUMAN | Q9v468 homo sapien |
| 21         | 42    | 27.3        | 402    | 1 EX7L_STRCO | O9fbm3 streptomyce |
| 22         | 42    | 27.3        | 926    | 1 UVRA_AQUAE | O66911 aquifex aeo |
| 23         | 42    | 27.3        | 986    | 1 CYGR_ARBPU | P1528 arabacia pun |
| 24         | 41.5  | 26.9        | 345    | 1 ARGC_BACHD | Q9K8v2 bacillus ha |
| 25         | 41.5  | 26.9        | 1882   | 1 Y468_MYCPN | P75109 mycoplasma  |
| 26         | 41.5  | 26.9        | 2200   | 1 LAR_CAEEL  | Q9bmn8 caenorhabdi |
| 27         | 41    | 26.6        | 162    | 1 PHA1_FREDI | P07122 fremyella d |
| 28         | 41    | 26.6        | 162    | 1 PHA1_PSEA9 | Q52447 pseudanaba  |
| 29         | 41    | 26.6        | 162    | 1 PHA2_FREDI | P06040 fremyella d |
| 30         | 41    | 26.6        | 162    | 1 PHCA_CYACA | O19910 cyanidium c |
| 31         | 41    | 26.6        | 162    | 1 PHCA_GALSU | P00306 galdieria s |
| 32         | 41    | 26.6        | 162    | 1 PHCA_SYNP2 | P03943 synechococc |
| 33         | 41    | 26.6        | 162    | 1 PHCA_SYNP3 | Q54715 synechocyst |

34 41 26.6 342 1 Y762 METJA Q58172 methanococc  
35 41 26.6 347 1 Y576 METJA Q57996 methanococc  
36 41 26.6 417 1 EX7L\_CORGL Q8nm3 corynebacte  
37 41 26.6 441 1 ZRAF\_ECOLI P14375 escherichia  
38 41 26.6 445 1 CLUS\_CANPA P25473 canis famil  
39 41 26.6 447 1 EX7L\_CORBF Q8fgp1 corynebacte  
40 41 26.6 510 1 YCGB\_ECOLI P29013 escherichia  
41 41 26.6 757 1 IMMT\_MOUSE Q8caq8 mus musculu  
42 41 26.6 798 1 UNR\_HUMAN O75534 homo sapien  
43 41 26.6 940 1 UVRA\_VIBCH Q9kuw5 vibrio chol  
44 41 26.6 973 1 UVRA\_RHILO Q98ma6 rhizobium l  
45 41 26.6 973 1 UVRA\_RHIME P56899 rhizobium m

#### ALIGNMENTS

RESULT 1  
PRRP RAT  
ID PRRP RAT STANDARD; PRT; 83 AA.  
AC P81278; Q8K3Y0;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DB Prolactin-releasing peptide precursor (PrRP) (Prolactin-releasing hormone) [Contains: Prolactin-releasing peptide PrRP31; Prolactin-releasing peptide PrRP20].  
GN PRH.  
OS Rattus norvegicus (Rat).  
OC Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RC TISSUE=Brain;  
RX MEDLINE=38268781; PubMed=9607765;  
RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S., Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M., Kurokawa T., Nishimura O., Onda H., Fujino M.;  
RA "A prolactin-releasing peptide in the brain."  
RT Nature 393:272-276(1998).  
RL [2]  
RN SEQUENCE FROM N.A. (ISOFORM 2).  
RC STRAIN=Sprague-Dawley; TISSUE=Hypothalamus;  
RA Anderson S.T., Kokay I.C., Lang T., Grattan D.R., Curlewis J.D.;  
RA "Quantitation of prolactin-releasing peptide (PrRP) mRNA expression in specific brain regions during the rat oestrous cycle and in lactation."  
RT Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP TISSUE SPECIFICITY.  
RX MEDLINE=99426652; PubMed=10498338;  
RA Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Habata Y., Hinuma S., Sekiguchi M., Kitada C., Kurokawa T., Nishimura O., Onda H., Sumino Y., Fujino M.;  
RA "Tissue distribution of prolactin-releasing peptide (PrRP) and its receptor."  
RT Regul. Pept. 83:1-10(1999).  
CC -1- FUNCTION: Stimulates prolactin (PRL) release and regulates the expression of prolactin through its receptor GPR10. May stimulate lactotrophs directly to secrete PRL.  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=1;  
CC IsoId=P81278-1; Sequence=Displayed;  
CC Name=2;  
CC IsoId=P81278-2; Sequence=VSP\_004370;  
CC -1- TISSUE SPECIFICITY: Widely expressed, with highest levels in medulla oblongata and hypothalamus.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its

FT MOD RES 53 AMIDATION (G-54 PROVIDE AMIDE GROUP) .  
SQ SEQUENCE 98 AA; 10544 MW; 08AC35A13B0FA908 CRC64;  
Query Match 94.8%; Score 146; DB 1; Length 98;  
Best Local Similarity 83.9%; Pred. No. 1.2e-16;  
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;





FT DOMAIN 610 670 CSD 8;  
SQ DOMAIN 674 735 CSD 9;  
SEQUENCE 798 AA; 88994 MW; F484B3FA8B0995A4 CRC64;

Query Match 29.9%; Score 46; DB 1; Length 798;  
Best Local Similarity 39.1%; Pred. No. 14;  
Matches 9; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 6 HSMEXRTDPDINPAWYXXRGIRPV 28  
||| : ||| : ||| :  
DB 583 HSVNGITEANPTIYGKVRPL 605

RESULT 6  
NER3\_BOVIN  
ID NER3\_BOVIN STANDARD; PRT; 428 AA.  
AC O97859;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DE Sialidase 3 (EC 3.2.1.18) (Membrane sialidase)  
DE (N-acetyl-alpha-neuraminidase 3).  
GN NEU3.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND CHARACTERIZATION.  
RC TISSUE=Brain;  
RX MEDLINE=99143165; PubMed=9988745;  
RA Miyagi T., Wada T., Iwamatsu A., Hata K., Yoshikawa Y., Tokuyama S., Sawada M.;  
RT "Molecular cloning and characterization of a plasma membrane-associated sialidase specific for gangliosides.";  
RL J. Biol. Chem. 274:5004-5011(1999).  
CC -!- FUNCTION: Plays a role in modulating the ganglioside content of the lipid bilayer at the level of membrane-bound sialyl glycoconjugates.  
CC -!- CATALYTIC ACTIVITY: Hydrolysis of alpha-(2->3)-, alpha-(2->6)-, alpha-(2->8)-glycosidic linkages of terminal sialic residues in oligosaccharides, glycoproteins, glycolipids, colominic acid and synthetic substrates.  
CC -!- SUBCELLULAR LOCATION: Membrane-associated (By similarity).  
CC -!- TISSUE SPECIFICITY: Expressed in brain.  
CC -!- SIMILARITY: BELONGS TO FAMILY 33 OF GLYCOSYL HYDROLASES.  
CC -!- SIMILARITY: Contains 3 BNR repeats.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.ebi.ac.uk/announcements or send an email to license@ebi.ac.uk).

EMBL; AB008184; BAA75071.1; --  
DR InterPro; IPR002860; GH\_BNR.  
DR Pfam; PF02012; BNR\_3.  
KW Hydrolase; Glycosidase; Membrane; Repeat.  
FT REPEAT 129 140 BNR 1.  
FT REPEAT 203 214 BNR 2.  
FT REPEAT 254 265 BNR 3.  
FT SITE 24 27 FRIP MOTIF.  
FT ACT\_SITE 25 25 By similarity.  
FT ACT\_SITE 45 45 Potential.  
FT ACT\_SITE 50 50 Potential.  
FT ACT\_SITE 87 87 Potential.  
FT ACT\_SITE 225 225 Potential.  
FT ACT\_SITE 245 245 Potential.  
FT ACT\_SITE 341 341 By similarity.  
FT ACT\_SITE 371 371 Potential.

FT ACT\_SITE 388 388 Potential  
SQ SEQUENCE 428 AA; 47916 MW; 418B34F3245A8P21 CRC64;

Query Match 29.2%; Score 45; DB 1; Length 428;  
Best Local Similarity 29.6%; Pred. No. 11;  
Matches 8; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

QY 2 RXHXSMEKRTPDINPAWYXXRGIRPV 28  
||| : ||| : ||| :  
DB 195 RAPHSLSMIYDDLGATWHGRLIKPM 221

RESULT 7  
PLSC\_HELPY  
ID PLSC\_HELPY STANDARD; PRT; 240 AA.  
AC O25903;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE 1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51) (1-AGPAT acyltransferase) (1-AGPAT) (lysophosphatidic acid acyltransferase) (LPAAT).  
GN PLSC OR HP1348.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
OC Helicobacteraceae; Helicobacter.  
NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=26695 / ATCC 700392;  
RX MEDLINE=97394467; PubMed=9252185;  
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G., Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A., Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S., Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A., McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K., Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M., Cotton M.D., Weidman J.M., Fujii C., Bowman C., Matthey L., Wallin B., Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M., Venter J.C.;  
RT "The complete genome sequence of the gastric pathogen Helicobacter pylori";  
RL Nature 388:539-547(1997).  
CC -!- FUNCTION: CONVERTS LYSPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION.  
CC -!- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate = CoA + 1,2-diacyl-sn-glycerol 3-phosphate.  
CC -!- PATHWAY: De novo phospholipid biosynthesis; second step.  
CC -!- SUBCELLULAR LOCATION: Inner membrane-associated (potential).  
CC -!- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.ebi.ac.uk/announcements or send an email to license@ebi.ac.uk).

EMBL; AB000636; AAD08393.1; --  
DR PIR; D64688; D64688.  
DR TIGR; HP1348; --  
DR InterPro; IPR002123; Acyltransferase.  
DR InterPro; IPR004552; AGP acyltrn.  
DR Pfam; PF01553; Acyltransferase; 1.  
DR SMART; SM00563; Plac; 1.  
DR TIGRFAMs; TIGR00530; AGP acyltrn; 1.  
KW Phospholipid biosynthesis; Transferase; Acyltransferase;  
KM Inner membrane; Complete proteome.  
SQ SEQUENCE 240 AA; 27745 MW; 22BD5D0EB190BBDD CRC64;

Query Match 28.9%; Score 44.5; DB1; Length 240;

```

Best Local Similarity 42.9%; Pred. No. 7;
Matches 9; Conservative 2; Mismatches 9; Indels 1; Gaps 1;

QY 1 SRXHXSMEXRTPDIN-PAWY 20
Db 197 ARTRLVLESYTPDFNSPTWY 217

RESULT 8
Y587 PASMU
ID Y587 PASMU STANDARD; PRT; 288 AA.
AC Q9CNS56;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein PM0587.
GN PM0587.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pm70;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70."
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465 (2001).
CC -1- SIMILARITY: BELONGS TO THE FRUCTOSAMINE KINASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB006094; AAK02671.1; -
DR InterPro; IPR005581; Fructosamin kin.
DR Pfam; PF03881; Fructosamin kin; 1.
KW Hypothetical protein; Kinase; Complete proteome.
SQ SEQUENCE 288 AA; 33778 MW; F4D2F6C26014D940 CRC64;

Query Match 28.6%; Score 44; DB 1; Length 288;
Best Local Similarity 34.8%; Pred. No. 10;
Matches 8; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 6 HSMEXRTPDINPAWYXRGIRPV 28
Db 21 HKKIHGTGEMHEAWIIDGQIPV 43

RESULT 9
NER3 HUMAN
ID NER3 HUMAN STANDARD; PRT; 428 AA.
AC Q9UQ49; Q9NQEI;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Sialidase 3 (EC 3.2.1.18) (Membrane sialidase) (Ganglioside sialidase)
DE (N-acetyl-alpha-neuraminidase 3).
GN NEU3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=99335353; PubMed=10405317;
RA Wada T., Yoshikawa Y., Tokuyama S., Kuwabara M., Akita H., Miyagi T.;
RT "Cloning, expression, and chromosomal mapping of a human ganglioside
sialidase."
Biochem. Biophys. Res. Commun. 261:21-27 (1999).
[2]
SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX PubMed=10861246;
RA Monti B., Bassi M.T., Papini N., Riboni M., Manzoni M., Venerando B.,
RA Croci G., Preti A., Ballabio A., Tettamanti G., Borsani G.;
RT "Identification and expression of NEU3, a novel human sialidase
associated to the plasma membrane."
Biochem. J. 349:343-351 (2000).
RL FUNCTION: Plays a role in modulating the ganglioside content of
the lipid bilayer at the level of membrane-bound sialyl
glycoconjugates.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of alpha-(2->3)-, alpha-(2->6)-,
CC alpha-(2->8)-glycosidic linkages of terminal sialic residues in
CC oligosaccharides, glycoproteins, glycolipids, colominic acid and
CC synthetic substrates.
CC -1- SUBCELLULAR LOCATION: Membrane-associated.
CC -1- TISSUE SPECIFICITY: Highly expressed in skeletal muscle, testis,
CC adrenal gland and thymus, followed by pancreas, liver, heart and
CC thymus. Weakly expressed in kidney, placenta, brain and lung.
CC -1- MISCELLANEOUS: Optimum pH is 3.8.
CC -1- SIMILARITY: BELONGS TO FAMILY 33 OF GLYCOSYL HYDROLASES.
CC -1- SIMILARITY: Contains 3 BNR repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB008185; BAA82611.1; -
DR EMBL; Y18563; CAB96131.1; ALT_INIT.
DR Genew; HGNC:7760; NEU3.
DR MIM; 604617; -
DR GO; GO:0005887; C: integral to plasma membrane; TAS.
DR GO; GO:0006689; P: ganglioside catabolism; TAS.
DR InterPro; IPR002860; GH_BNR.
DR Pfam; PF02012; BNR; 3.
KW Hydrolase; Glycosidase; Membrane; Repeat.
FT REPEAT 129 140 BNR 1.
FT REPEAT 203 214 BNR 2.
FT REPEAT 254 265 BNR 3.
FT SITE 24 27 FRIP MOTIF.
FT ACT_SITE 25 25 By similarity.
FT ACT_SITE 45 45 Potential.
FT ACT_SITE 50 50 Potential.
FT ACT_SITE 87 87 Potential.
FT ACT_SITE 225 225 Potential.
FT ACT_SITE 245 245 Potential.
FT ACT_SITE 340 340 By similarity.
FT ACT_SITE 370 370 Potential.
FT ACT_SITE 387 387 Potential.
SQ SEQUENCE 428 AA; 48252 MW; 35D1DD9359A78C98 CRC64;

Query Match 28.6%; Score 44; DB 1; Length 428;
Best Local Similarity 34.8%; Pred. No. 15;
Matches 8; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 6 HSMEXRTPDINPAWYXRGIRPV 28
Db 199 HSLMIYSDDLGVTHHGRIRPV 221

RESULT 10
UVRA METTH
ID UVRA METTH STANDARD; PRT; 962 AA.
AC Q26543;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)

```

DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE UvrABC system protein A (UvrA protein) (Excinuclease ABC subunit A).  
GN UVR A OR MTH443.  
OS Methanobacterium thermoautotrophicum.  
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;  
OC Methanobacteriaceae; Methanothermobacter.  
OX NCBI\_TaxID=187420;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Deita H;  
RX MEDLINE=98037514; PubMed=9371463;  
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,  
RA Alldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,  
RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,  
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,  
RA Jiwani N., Caruso A., Bush D., Safer H., Fatwell D., Prabhakar S.,  
RA McDougall S., Shiner G., Goyal A., Pietrowski S., Church G.M.,  
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;  
RT "Complete genome sequence of Methanobacterium thermoautotrophicum  
deltaH: functional analysis and comparative genomics.";  
RL J. Bacteriol. 179:7135-7155(1997).  
CC -1- FUNCTION: The UvrABC repair system catalyzes the recognition and  
CC processing of DNA lesions. UvrA is an ATPase and a DNA-binding  
CC protein. A damage recognition complex composed of 2 uvrA and 2  
CC uvrB subunits scans DNA for abnormalities. When the presence of a  
CC lesion has been verified by uvrB, the uvrA molecules dissociate  
CC (by similarity).  
CC -1- SUBUNIT: Forms a heterotetramer with uvrB during the search for  
CC lesions (by similarity).  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. UVR A SUBFAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC  
CC EMBL; AE000828; AAB84949.1; -  
CC PIR; H69157; H69157.  
CC HAMAP; MF 00205; -1.  
CC InterPro; IPR003593; AAA\_ATPase.  
CC InterPro; IPR003439; ABC\_transporter.  
CC InterPro; IPR004602; UvrA.  
CC Pfam; PF00005; ABC\_tran; 2.  
CC ProDom; PD000006; ABC\_transporter; 1.  
CC SMART; SM00382; AAA; 2.  
CC TIGRfam; TIGR00630; UvrA; 1.  
CC PROSITE; PS00211; ABC\_TRANSPORTER\_1; 2.  
CC PROSITE; PS00893; ABC\_TRANSPORTER\_2; 2.  
CC SOS response; Excision nuclease; DNA repair; DNA recombination;  
CC DNA excision; ATP-binding; DNA-binding; Repeat; Zinc; Metal-binding;  
CC Zinc-finger; Complete proteome.  
CC FT BIND 38 45 ATP (POTENTIAL).  
CC FT NP\_BIND 649 656 ATP (POTENTIAL).  
CC FT ZN\_BIND 748 774 C4-TYPE.  
CC SQ SEQUENCE 962 AA; 108395 MW; 2C0EF7FC41CCD060 CRC64;  
  
Query Match 28.6%; Score 44; DB 1; Length 962;  
Best Local Similarity 36.8%; Pred. No. 35;  
Matches 14; Conservative 0; Mismatches 6; Indels 18; Gaps 2;  
  
Qy 11 RTDINPAWYX-----XRGIRPVGRF 31  
||| |||  
Db 703 RTRSNPATYGVTFTHIRLFAQTPEARKGYRP-GRF 739  
||| |||  
  
RESULT 11  
OPSI LIMPO  
ID \_OPSI LIMPO PRT; 376 AA.  
AC P35360;

DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Lateral eye opsin.  
OS Limulus polyphemus (Atlantic horseshoe crab).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Merostomata; Xiphosura;  
OC Limulidae; Limulus.  
OX NCBI\_TaxID=6850;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lateral eye;  
RX MEDLINE=99317641; PubMed=8327495;  
RA Smith W.C., Price D.A., Greenberg R.M., Battelle B.-A.;  
RT "Opsins from the lateral eyes and ocelli of the horseshoe crab,  
RL Limulus polyphemus.";  
RL Proc. Natl. Acad. Sci. U.S.A. 90:6150-6154(1993).  
CC -1- FUNCTION: VISUAL PIGMENTS ARE THE LIGHT-ABSORBING MOLECULES THAT  
CC MEDIATE VISION. THEY CONSIST OF AN APOPROTEIN, OPSIN, COVALENTLY  
CC LINKED TO CIS-RETINAL.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -1- TISSUE SPECIFICITY: LATERAL EYE.  
CC -1- PTM: SOME OR ALL OF THE CARBOXYL-TERMINAL SER OR THR RESIDUES MAY  
CC BE PHOSPHORYLATED (BY SIMILARITY).  
CC -1- MISCELLANEOUS: THIS OPSIN HAS AN ABSORPTION MAXIMUM AT 520 NM.  
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
CC OPSIN SUBFAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC  
CC EMBL; L03791; AAA28273.1; -  
CC EMBL; L03781; AAA02498.1; -  
CC PIR; B48197; B48197.  
CC HSSP; P02699; LEDV.  
CC InterPro; IPR000276; GPCR\_Rhodopsn.  
CC InterPro; IPR001760; Opsin.  
CC Pfam; PF00001; 7tm\_1; 1.  
CC PRINTS; PR00237; GPCRHHODOPSN.  
CC PROSITE; PS00237; G\_PROTEIN\_RECP\_F1\_1; 1.  
CC PROSITE; PS00362; G\_PROTEIN\_RECP\_F1\_2; 1.  
CC PROSITE; PS00238; OPSIN; 1.  
CC KW Photoreceptor; Retinal protein; Transmembrane; Glycoprotein; Vision;  
CC Phosphorylation; G-protein coupled receptor.  
CC FT DOMAIN 1 46  
CC FT TRANSMEM 47 71 1 (POTENTIAL).  
CC FT DOMAIN 72 83 CYTOPLASMIC.  
CC FT TRANSMEM 84 108 2 (POTENTIAL).  
CC FT DOMAIN 109 123 3 (POTENTIAL).  
CC FT TRANSMEM 124 143 3 (POTENTIAL).  
CC FT DOMAIN 144 162 CYTOPLASMIC.  
CC FT TRANSMEM 163 186 4 (POTENTIAL).  
CC FT DOMAIN 187 210 5 (POTENTIAL).  
CC FT TRANSMEM 211 238 6 (POTENTIAL).  
CC FT DOMAIN 239 274 6 (POTENTIAL).  
CC FT TRANSMEM 275 298 7 (POTENTIAL).  
CC FT DOMAIN 299 306 7 (POTENTIAL).  
CC FT TRANSMEM 307 331 7 (POTENTIAL).  
CC FT DOMAIN 332 376 CYTOPLASMIC.  
CC FT DISULFID 120 137 BY SIMILARITY.  
CC FT BINDING 318 318 RETINAL CHROMOPHORE (BY SIMILARITY).  
CC FT CARBOHYD 17 17 N-LINKED (GLCNAC. .) (POTENTIAL).  
CC FT CARBOHYD 193 193 N-LINKED (GLCNAC. .) (POTENTIAL).  
CC SQ SEQUENCE 376 AA; 42139 MW; CCB401766AB06F26 CRC64;  
  
Query Match 27.9%; Score 43; DB 1; Length 376;  
Best Local Similarity 46.2%; Pred. No. 19;  
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

```
Oy 13 PDINPAWXXRGI 25
Db 40 PPMNPLWYSILGV 52

RESULT 12
OPS2 LIMPO
ID OPS2 LIMPO STANDARD; PRT; 376 AA.
AC P35361;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ocellar opsin.
OS Limulus polyphemus (Atlantic horseshoe crab).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Merostomata; Xiphosura;
OC Limulidae; Limulus.
OX NCBI_TaxID=6850;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Median ocelli;
RX MEDLINE=93317641; PubMed=8327495;
RA Smith W.C., Price D.A., Greenberg R.M., Battelle B.-A.;
RT "Opsins from the lateral eyes and ocelli of the horseshoe crab,
RL Limulus polyphemus."
RL Proc. Natl. Acad. Sci. U.S.A. 90:6150-6154(1993).
CC -1- FUNCTION: VISUAL PIGMENTS ARE THE LIGHT-ABSORBING MOLECULES THAT
CC MEDIATE VISION. THEY CONSIST OF AN APOPROTEIN, OPSIN, COVALENTLY
CC LINKED TO CIS-RETINAL.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: OCELLAR CELLS; MEDIAN OCELLI.
CC -1- PTM: SOME OR ALL OF THE CARBOXYL-TERMINAL SER OR THR RESIDUES MAY
CC BE PHOSPHORYLATED (BY SIMILARITY).
CC -1- MISCELLANEOUS: THIS OPSIN HAS AN ABSORPTION MAXIMUM AT 530 NM.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC OPSIN SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L03792; AAA28274.1; -.
CC EMBL; L03782; AAA02499.1; -.
CC PIR; A48197; A48197.
CC HSSP; P02699; 1EDV.
CC InterPro; IPR000276; GPCR_Rhodopsn.
CC DR Pfam; PF00001; 7tm_1; 1.
CC PRINTS; PR00237; GPCRHHODOPSN.
CC PROSITE; PS00237; G-PROTEIN RECEPTOR_F1_1; 1.
CC PROSITE; PS00262; G-PROTEIN RECEPTOR_F1_2; 1.
CC PROSITE; PS00238; OPSIN; 1.
CC Photoreceptor; Retinal protein; Transmembrane; Glycoprotein; Vision;
KW Phosphorylation; G-protein coupled receptor.
FT DOMAIN 1 46
FT TRANSFAM 47 71
FT DOMAIN 72 83
FT TRANSFAM 84 108
FT DOMAIN 109 123
FT TRANSFAM 124 143
FT DOMAIN 144 162
FT TRANSFAM 163 186
FT DOMAIN 187 210
FT TRANSFAM 211 238
FT DOMAIN 239 274
FT TRANSFAM 275 298
FT DOMAIN 299 306
FT TRANSFAM 307 331
FT DOMAIN 332 376
FT TRANSFAM 377 399
FT DISULFID 120 197

Oy 13 PDINPAWXXRGI 25
Db 40 PPMNPLWYSILGV 52

Query Match 27.9%; Score 43; DB 1; Length 376;
Best Local Similarity 46.2%; Pred. No. 19;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Oy 13 PDINPAWXXRGI 25
Db 40 PPMNPLWYSILGV 52

RESULT 13
CYCR_CHRVI
ID CYCR_CHRVI STANDARD; PRT; 383 AA.
AC O82947;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Photosynthetic reaction center cytochrome C subunit precursor.
GN PUPC.
OS Chromatium vinosum.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Chromatiales;
OC Chromatiaceae; Allochromatium.
OX NCBI_TaxID=1049;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D / ATCC 17899 / DSM 180;
RA Corson G.B., Nagashima K.V., Mateuura K., Sakuragi Y., Ruwanthi W.,
RA Qin H., Allen R., Knaff D.B.;
RT "Primary structure of genes encoding light-harvesting and reaction
RL center proteins from Chromatium vinosum."
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THE REACTION CENTER OF PURPLE BACTERIA CONTAIN A
CC TIGHTLY BOUND CYTOCHROME MOLECULE WHICH REDUCES THE PHOTO
CC OXIDIZED PRIMARY ELECTRON DONOR.
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
CC (By similarity).
CC -1- PTM: BINDS FOUR HEME GROUPS PER MOLECULE.
CC -1- SIMILARITY: HIGH, WITH OTHER PHOTOSYNTHETIC REACTION CENTER
CC CYTOCHROME C SUBUNITS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB011811; BAA32742.1; -.
CC HSSP; P07173; 6PRC.
CC InterPro; IPR000345; CytC_heme_bind.
CC InterPro; IPR003158; CytC_RC.
CC Pfam; PF02276; CytC_RC; 1.
CC ProDom; PD010011; CytC_RC; 1.
CC PROSITE; PS00190; CYTOCHROME C; 4.
CC Electron transport; Photosynthesis; Reaction center; Heme;
KW Membrane; Lipoprotein; Repeat; Signal.
FT SIGNAL 1 22
FT CHAIN 23 383
FT BY SIMILARITY.
FT PHOTOSYNTHETIC REACTION CENTER CYTOCHROME
FT C SUBUNIT.
FT LIPID 23 23
FT METAL 94 94
FT N-ACYL DIGLYCERIDE (BY SIMILARITY).
FT IRON (HEME 1 AXIAL LIGAND)
FT (BY SIMILARITY).
FT BINDING 107 107
FT HEME 1 (COVALENT) (BY SIMILARITY).
FT BINDING 110 110
FT HEME 1 (COVALENT) (BY SIMILARITY).
FT METAL 111 111
FT IRON (HEME 1 AXIAL LIGAND)
FT (BY SIMILARITY).
FT METAL 130 130
FT IRON (HEME 2 AXIAL LIGAND)
FT (BY SIMILARITY).
```

```
FT METAL 144 144 IRON (HEME 4 AXIAL LIGAND)
FT (BY SIMILARITY).
FT BINDING 152 152 HEME 2 (COVALENT) (BY SIMILARITY).
FT BINDING 155 155 HEME 2 (COVALENT) (BY SIMILARITY).
FT METAL 156 156 IRON (HEME 2 AXIAL LIGAND)
FT (BY SIMILARITY).
FT METAL 236 236 IRON (HEME 3 AXIAL LIGAND)
FT (BY SIMILARITY).
FT BINDING 247 247 HEME 3 (COVALENT) (BY SIMILARITY).
FT BINDING 250 250 HEME 3 (COVALENT) (BY SIMILARITY).
FT METAL 251 251 IRON (HEME 3 AXIAL LIGAND)
FT (BY SIMILARITY).
FT BINDING 307 307 HEME 4 (COVALENT) (BY SIMILARITY).
FT BINDING 310 310 HEME 4 (COVALENT) (BY SIMILARITY).
FT METAL 311 311 IRON (HEME 4 AXIAL LIGAND)
FT (BY SIMILARITY).
SQ SEQUENCE 383 AA; 41522 MW; 96BCD91FF1B9AE7B CRC64;

Query Match 27.9%; Score 43; DB 1; Length 383;
Best Local Similarity 36.8%; Pred. No. 20;
Matches 7; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 12 TPDINPWTXXRGIRPVGR 30
DB 263 TPQRTTAWYAIRVRDQ 281

RESULT 14
YAAJ_ECOLI
ID YAAJ_ECOLI STANDARD; PRT; 476 AA.
AC P30143;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative transporter yaaJ.
GN YAAJ OR B0007.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
ON NCBI_TaxID=562;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=92334977; PubMed=1630901;
RA Yura T., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N.,
RA Isono K., Mizobuchi K., Nakata A.;
RT "Systematic sequencing of the Escherichia coli genome: analysis of
RT the 0-2.4 min region."
RL Nucleic Acids Res. 20:3305-3308(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MGL655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Ferna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (potential).
CC -1- SIMILARITY: BELONGS TO THE SODIUM-ALANINE SYMPORTER FAMILY
CC (SAP). STRONG, TO H.INFLUENZAE HI0183.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D10483; BAB96585.1; --

DR YAAJ_ECOLI
DR PIR; A42743; A42743.
DR HSSP; P29253; 1RIL.

DR METAL 144 144 IRON (HEME 4 AXIAL LIGAND)
DR (BY SIMILARITY).
DR BINDING 152 152 HEME 2 (COVALENT) (BY SIMILARITY).
DR BINDING 155 155 HEME 2 (COVALENT) (BY SIMILARITY).
DR METAL 156 156 IRON (HEME 2 AXIAL LIGAND)
DR (BY SIMILARITY).
DR METAL 236 236 IRON (HEME 3 AXIAL LIGAND)
DR (BY SIMILARITY).
DR BINDING 247 247 HEME 3 (COVALENT) (BY SIMILARITY).
DR BINDING 250 250 HEME 3 (COVALENT) (BY SIMILARITY).
DR METAL 251 251 IRON (HEME 3 AXIAL LIGAND)
DR (BY SIMILARITY).
DR BINDING 307 307 HEME 4 (COVALENT) (BY SIMILARITY).
DR BINDING 310 310 HEME 4 (COVALENT) (BY SIMILARITY).
DR METAL 311 311 IRON (HEME 4 AXIAL LIGAND)
DR (BY SIMILARITY).
SQ SEQUENCE 383 AA; 41522 MW; 96BCD91FF1B9AE7B CRC64;

Query Match 27.9%; Score 43; DB 1; Length 476;
Best Local Similarity 44.4%; Pred. No. 25;
Matches 8; Conservative 2; Mismatches 2; Indels 6; Gaps 1;

QY 14 DIN-----PAWYXXRGI 25
DB 120 DVNGQFRGGPAWYMARGL 137

RESULT 15
POL_MLVK
ID POL_MLVK STANDARD; PRT; 581 AA.
AC P31795;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Pol polyprotein [Contains: Protease (EC 3.4.23.-); Reverse
DE transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)] (Fragment).
GN POL.
OS Radiation murine leukemia virus (strain Kaplan).
OC Viruses; Retroviridae; Retroviridae; Gammaretrovirus.
ON NCBI_TaxID=31689;
RX [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=92333703; PubMed=1629969;
RA Poliquin L., Bergeron D., Fortier J.L., Paquette Y., Bergeron R.,
RA Rassart E.;
RT "Determinants of thymotropism in Kaplan radiation leukemia virus and
RT nucleotide sequence of its envelope region."
RL J. Virol. 66:5141-5146(1992).
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + [DNA] (N).
CC -1- PTM: SPECIFIC ENZYMIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A GAG-POL
CC POLYPROTEIN.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A2.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M93052; AAA46525.1; --
CC PIR; A42743; A42743.
CC HSSP; P29253; 1RIL.
```

DR MEROPS: A02.008; --  
DR InterPro: IPR001969; Aspprotease\_site.  
DR InterPro: IPR002156; RNaseH.  
DR InterPro: IPR001584; Rve.  
DR Pfam: PF00075; rnaseH; 1.  
DR PROSITE: PS00141; ASP\_PROTEASE; PARTIAL.  
KW Hydrolase; Transferase; RNA-directed DNA polymerase;  
KW Aspartyl protease; Endonuclease; Polypeptidase.  
FT NON\_TER 1  
SQ SEQUENCE 581 AA; 65157 MW; 8D7A38694C8E036E CRC64;  
Query Match 27.9%; Score 43; DB 1; Length 581;  
Best Local Similarity 34.8%; Pred. No. 30;  
Matches 8; Conservative 2; Mismatches 13; Indels 0; Gaps 0;  
QY 6 HSMEXRTPDINPWTYXRGIRPV 28  
Db 550 HVKAATTPPIRPSWRVQRSONEL 572

Search completed: December 3, 2003, 19:12:03  
Job time : 6.25 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: December 3, 2003, 19:03:25 ; Search time 26.5 Seconds  
(without alignments)  
301.873 Million cell updates/sec

Title: US-09-868-885B-44

Perfect score: 154

Sequence: 1 SRXHXSMEXRTPDINPAWYXXRGIRPVGRF 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

1: SPREMBL 23:\*

1: sp\_archaea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mhc:\*

8: sp\_organelle:\*

9: sp\_phase:\*

10: sp\_plant:\*

11: sp\_rodent:\*

12: sp\_virus:\*

13: sp\_vertebrate:\*

14: sp\_unclassified:\*

15: sp\_virus:\*

16: sp\_bacteriap:\*

17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID     | Description        |
|------------|-------|-------------|--------|-----------|--------------------|
| 1          | 146   | 94.8        | 98     | 6 Q8WN12  | Q8WN12 ovis aries  |
| 2          | 98    | 63.6        | 117    | 13 Q9W624 | Q9W624 carassius a |
| 3          | 50    | 32.5        | 692    | 2 Q18J6   | Q18J6 rhodospiril  |
| 4          | 49.5  | 32.1        | 664    | 16 Q91W4  | Q91W4 pseudomonas  |
| 5          | 49    | 31.8        | 220    | 16 Q91022 | Q91022 pseudomonas |
| 6          | 49    | 31.8        | 790    | 10 Q9M371 | Q9M371 arabidopsis |
| 7          | 47    | 30.5        | 128    | 16 P74747 | P74747 synchocyst  |
| 8          | 47    | 30.5        | 173    | 17 Q97A00 | Q97A00 thermoplasm |
| 9          | 47    | 30.5        | 1253   | 2 Q8RTV7  | Q8RTV7 uncultured  |
| 10         | 47    | 30.5        | 1271   | 2 Q91SV4  | Q91SV4 chlorobium  |
| 11         | 47    | 30.5        | 1278   | 16 Q8K339 | Q8K339 chlorobium  |
| 12         | 46.5  | 30.2        | 416    | 16 Q8K746 | Q8K746 streptococc |
| 13         | 46.5  | 30.2        | 420    | 2 Q93L27  | Q93L27 streptomyce |
| 14         | 46    | 29.9        | 54     | 4 Q9UJF9  | Q9UJF9 homo sapien |
| 15         | 46    | 29.9        | 118    | 16 Q8UKM3 | Q8UKM3 agrobacteri |
| 16         | 46    | 29.9        | 314    | 16 O06348 | O06348 mycobacteri |

|    |      |      |      |           |                    |
|----|------|------|------|-----------|--------------------|
| 17 | 46   | 29.9 | 333  | 16 Q9PH76 | Q9PH76 xylella fas |
| 18 | 46   | 29.9 | 348  | 5 O45431  | O45431 caenorhabdi |
| 19 | 46   | 29.9 | 450  | 11 Q9DA19 | Q9DA19 mus musculu |
| 20 | 46   | 29.9 | 465  | 4 O60687  | O60687 homo sapien |
| 21 | 46   | 29.9 | 465  | 4 Q8W85   | Q8W85 homo sapien  |
| 22 | 46   | 29.9 | 540  | 10 Q9LGZ0 | Q9LGZ0 oryza sativ |
| 23 | 46   | 29.9 | 689  | 11 Q8R3R1 | Q8R3R1 mus musculu |
| 24 | 46   | 29.9 | 767  | 11 Q8JZN2 | Q8JZN2 mus musculu |
| 25 | 46   | 29.9 | 798  | 11 Q91W50 | Q91W50 mus musculu |
| 26 | 46   | 29.9 | 1236 | 2 Q9JPA4  | Q9JPA4 rhodocyclu  |
| 27 | 46   | 29.9 | 1245 | 2 Q8KZ57  | Q8KZ57 uncultured  |
| 28 | 46   | 29.9 | 1292 | 2 Q9ZG85  | Q9ZG85 hellobacill |
| 29 | 45.5 | 29.5 | 416  | 16 Q8P0K1 | Q8P0K1 streptococc |
| 30 | 45   | 29.2 | 240  | 16 Q9RSN5 | Q9RSN5 deinococcu  |
| 31 | 45   | 29.2 | 412  | 10 Q9FVN0 | Q9FVN0 arabidopsis |
| 32 | 45   | 29.2 | 501  | 10 Q9LYX0 | Q9LYX0 arabidopsis |
| 33 | 45   | 29.2 | 637  | 16 Q9A5E9 | Q9A5E9 caulobacter |
| 34 | 45   | 29.2 | 767  | 5 Q20170  | Q20170 caenorhabdi |
| 35 | 45   | 29.2 | 813  | 17 Q8ZXH2 | Q8ZXH2 pyrobaculum |
| 36 | 45   | 29.2 | 1654 | 16 Q8FM04 | Q8FM04 corynebacte |
| 37 | 44.5 | 28.9 | 258  | 11 Q8BJ19 | Q8BJ19 mus musculu |
| 38 | 44.5 | 28.9 | 273  | 11 Q8S5B6 | Q8S5B6 mus musculu |
| 39 | 44.5 | 28.9 | 954  | 16 Q9KGP0 | Q9KGP0 bacillus ha |
| 40 | 44   | 28.6 | 72   | 16 Q9EYB3 | Q9EYB3 escherichia |
| 41 | 44   | 28.6 | 145  | 5 Q95Z85  | Q95Z85 leishmania  |
| 42 | 44   | 28.6 | 159  | 16 Q9PD41 | Q9PD41 xylella fas |
| 43 | 44   | 28.6 | 226  | 16 Q8RF10 | Q8RF10 fusbacteri  |
| 44 | 44   | 28.6 | 250  | 16 Q912P2 | Q912P2 pseudomonas |
| 45 | 44   | 28.6 | 284  | 17 O50128 | O50128 pyrococcus  |

## ALIGNMENTS

### RESULT 1

Q8WN12 ID Q8WN12 PRELIMINARY; PRT; 98 AA.

AC Q8WN12: 01-MAR-2002 (T-EMBLrel. 20, Created)

DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)

DT 01-MAR-2002 (T-EMBLrel. 20, Last annotation update)

DB Preprolactin-releasing peptide.

OS Ovis aries (Sheep).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Caprinae; Ovis.

OC NCBI\_TaxID=9940;

RN [1]

RP SEQUENCE FROM N.A.

RA Curlewis J.D., Kusters D.H.L., Barclay J.L., Anderson S.T.;

RT "Prolactin-releasing peptide (PRP) in the ewe: cDNA cloning, mRNA

RT distribution and effects on prolactin secretion in vitro and in

RT vivo.";

RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF450453; AAL47178.1; -

SQ SEQUENCE 98 AA; 10513 MW; 2A53331ED62CAAB5 CRC64;

Query Match 94.8%; Score 146; DB 6; Length 98;

Best Local Similarity 83.9%; Pred. No. 7.8e-16;

Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 SRXHXSMEXRTPDINPAWYXXRGIRPVGRF 31

Db 23 SRAHQSHMEIRTPDINPAWYAGRGIRPVGRF 53

### RESULT 2

Q9W624 ID Q9W624 PRELIMINARY; PRT; 117 AA.

AC Q9W624:

DT 01-NOV-1999 (T-EMBLrel. 12, Created)

DT 01-NOV-1999 (T-EMBLrel. 12, Last sequence update)

DT 01-DEC-2001 (T-EMBLrel. 19, Last annotation update)



[illegible]

## RESULT 6

Q9M371 ID Q9M371 PRELIMINARY; PRT; 790 AA.  
AC Q9M371;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE Hypothetical 87.4 kDa protein.  
GN P15G16.60  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eucosids II; Brassicales; Brassicaceae; Arabidopsiis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA De Haan M., Maare A.C., Grivell L.A., Mewes H.W., Lemcke K.,  
RA Mayer K.F.X., Quetier F., Salanoubat M.;  
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA EU Arabidopsis sequencing project;  
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL132959; CAB71097.1; -  
KW Hypothetical protein.  
SQ SEQUENCE 790 AA; 87376 MW; B222724B75690F30 CRC64;

Query Match 31.8%; Score 49; DB 10; Length 790;  
Best Local Similarity 47.4%; Pred. No. 32;  
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 13 PDINPAWYXXRGIRPVGRF 31  
Db 366 PFHPNRTYXSRGLQPHGRW 384

## RESULT 7

P74747 ID P74747 PRELIMINARY; PRT; 128 AA.  
AC P74747;  
DT 01-FEB-1997 (TREMBlrel. 02, Created)  
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)  
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
DE Hypothetical protein sir0601.  
GN SLR0601.  
OS Synechocystis sp. (strain PCC 6803).  
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.  
OX NCBI\_TaxID=1148;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97061201; PubMed=8905231;  
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,  
RA Miyajima N., Hiroseawa M., Sugitara M., Sasamoto S., Kimura T.,  
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,  
RA Shampo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,  
RA Tabata S.;  
RL "Sequence analysis of the genome of the unicellular cyanobacterium  
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the  
RT entire genome and assignment of potential protein-coding regions.";  
RL DNA Res. 3:109-136(1996).  
DR EMBL; D90317; BAA18867.1; -  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 128 AA; 13481 MW; 4FF5B364C38396F6 CRC64;

Query Match 30.5%; Score 47; DB 16; Length 128;  
Best Local Similarity 33.3%; Pred. No. 9.3;  
Matches 11; Conservative 5; Mismatches 13; Indels 4; Gaps 1;

Qy 1 SRXHXHSMEXRTPDINPAWYXXR-----GIRPVG 29  
Db 11 SPENNQLDCSSPDSRKWYQARLFSFGITPTG 43

## RESULT 8

Q97A00 ID Q97A00 PRELIMINARY; PRT; 173 AA.  
AC Q97A00;  
DT 01-OCT-2001 (TREMBlrel. 18, Created)  
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)  
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
DE Ferricytochrome b5 binding protein.  
GN TV1010 OR TVG1032219.  
OS Thermoplasma volcanum.  
OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;  
OC Thermoplasmataceae; Thermoplasma.  
OX NCBI\_TaxID=50339;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=GSS1 / DSM 4299 / JCM 9571;  
RX MEDLINE=20570466; PubMed=11121031;  
RA Kawashima T., Amano N., Koike H., Makino S.-I., Higuchi S.,  
RA Kawashima-Ohya Y., Watanabe K., Yamazaki M., Kanehori K., Kawamoto T.,  
RA Nunohiba T., Yamamoto Y., Aramaki H., Makino K., Suzuki M.;  
RT "Archaeal adaptation to higher temperatures revealed by genomic  
RT sequence of Thermoplasma volcanum.";  
RL Proc. Natl. Acad. Sci. U.S.A. 97:14257-14262(2000).  
DR EMBL; AP000995; BAB60152.1; -  
DR InterPro; IPR001451; Hexapep\_transf.  
DR Pfam; PF00132; hexapep; 4.  
KW Complete proteome.  
SQ SEQUENCE 173 AA; 19041 MW; 1A750DD8B21C7524 CRC64;

Query Match 30.5%; Score 47; DB 17; Length 173;  
Best Local Similarity 32.0%; Pred. No. 13;  
Matches 8; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

Qy 6 HSMEXRTPDINPAWYXXRGIRPVGR 30  
Db 3 YEPGVRPDIIDPSAYVSESATVIGK 27

## RESULT 9

Q8RTV7 ID Q8RTV7 PRELIMINARY; PRT; 1253 AA.  
AC Q8RTV7;  
DT 01-JUN-2002 (TREMBlrel. 21, Created)  
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)  
DE CoBN/magnesium chelatase family protein.  
GN ERAC000-65D09.24.  
OS uncultured proteobacterium.  
OC Bacteria; Proteobacteria; environmental samples.  
OX NCBI\_TaxID=153809;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21822632; PubMed=11832943;  
RA Beja O., Suzuki M.T., Heidelberg J.F., Nelson W.C., Preston C.M.,  
RA Hanada T., Eisen J.A., Fraser C.M., DeLong E.E.;  
RT "Unsuspected diversity among marine aerobic anoxygenic phototrophs.";  
RL Nature 415:630-633(2002).  
DR EMBL; AB008919; AAL76369.1; -  
DR InterPro; IPR003672; CoBN/Mg chltase.  
DR Pfam; PF02514; coBN-Mg\_chel; 1.  
SQ SEQUENCE 1253 AA; 137586 MW; C98ABC010E261511 CRC64;

Query Match 30.5%; Score 47; DB 2; Length 1253;  
Best Local Similarity 32.1%; Pred. No. 1.1e+02;  
Matches 9; Conservative 5; Mismatches 8; Indels 6; Gaps 1;

Qy 7 SMEXRTPDINPAWYXX-----RGIRPV 28  
Db 1133 ALETRTRVLPNPKWYESMLDHGEGVRAI 1160

## RESULT 10

Q93SV4

```

ID Q93SV4 PRELIMINARY; PRT; 1271 AA.
AC Q93SV4;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE BCH1.
GN Chlorobium tepidum.
OS Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;
OC Chlorobium.
OX NCBI_TaxID=1097;
RN [1]
RP MEDLINE=20433268; PubMed=10976061;
RA Xiong J., Fischer W.M., Inoue K., Nakahara M., Bauer C.B.;
RT "Molecular evidence for the early evolution of photosynthesis.";
RL Science 289:1724-1730(2000).
DR EMBL; AY005136; AAG12412.1; -
DR InterPro; IPR003672; COB/Mg chlase.
DR Pfam; PF02514; cobN-Mg chl; 1.
SQ SEQUENCE 1271 AA; 141905 MW; 752C9D216549A2B1 CRC64;

Query Match 30.5%; Score 47; DB 2; Length 1271;
Best Local Similarity 32.1%; Pred. No. 1.1e+02;
Matches 9; Conservative 5; Mismatches 8; Indels 6; Gaps 1;

QY 7 SMEXRTPDINPAWYXX-----RGIRPV 28
DB 1152 ALEARTKLNPKWYEAIEHGYGVREI 1179

RESULT 11
Q8KB39 PRELIMINARY; PRT; 1278 AA.
AC Q8KB39;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Magnesium-protoporphyrin methyltransferase.
GN BCHH-1 OR CT1957.
OS Chlorobium tepidum.
OC Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;
OC Chlorobium.
OX NCBI_TaxID=1097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TLS / ATCC 49652 / DSM 12025;
RX MEDLINE=22103685; PubMed=12093901;
RA Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.P., Wu M.,
RA Dodson R.J., Deboy R., Gwinn M.L., Nelson W.C., Haft D.H.,
RA Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.B., Yang F.,
RA Holt I., Unayam L.A., Mason T., Brenner M., Shea T.P., Parksey D.,
RA Nierman W.C., Feldblyum T.V., Hansen C.L., Craven M.B., Radune D.,
RA Vamathavan J., Khouri H., White O., Gruber T.M., Ketchum K.A.,
RA Venter J.C., Tettelin H., Bryant D.A., Fraser C.M.;
RT "The complete genome sequence of Chlorobium tepidum TLS, a
RT photosynthetic, anaerobic, green-sulfur bacterium.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002).
DR EMBL; A2012946; AAM73176.1; -
DR TIGR; CT1957; -
DR InterPro; IPR003672; COB/Mg chlase.
DR Pfam; PF02514; cobN-Mg chl; 1.
KW Transferase; Methyltransferase; Complete proteome.
SQ SEQUENCE 1278 AA; 142814 MW; 7983FE019A52AE22 CRC64;

Query Match 30.5%; Score 47; DB 16; Length 1278;
Best Local Similarity 32.1%; Pred. No. 1.1e+02;
Matches 9; Conservative 5; Mismatches 8; Indels 6; Gaps 1;

QY 7 SMEXRTPDINPAWYXX-----RGIRPV 28
DB 1159 ALEARTKLNPKWYEAIEHGYGVREI 1186

```

```

RESULT 12
Q8K746 PRELIMINARY; PRT; 416 AA.
AC Q8K746;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Putative secreted protein DltD.
GN DLTd OR SPY3_0991.
OS Streptococcus pyogenes (serotype M3).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OX NCBI_TaxID=198466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS315 / Serotype M3;
RX MEDLINE=22133808; PubMed=12122206;
RA Betes S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,
RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.P., Parkins L.D.,
RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA Schlievert P.M., Musher J.M.;
RT "Genome sequence of a serotype M3 strain of group A Streptococcus:
RT phage-encoded toxins, the high-virulence phenotype, and clone
RT emergence.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).
DR EMBL; AE014156; AAM79598.1; -
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 416 AA; 48086 MW; 87671525E013E45A CRC64;

Query Match 30.2%; Score 46.5; DB 16; Length 416;
Best Local Similarity 25.0%; Pred. No. 40;
Matches 9; Conservative 7; Mismatches 11; Indels 9; Gaps 1;

QY 1 SRXHXSMEXRTPD-----INPAWYXXRGIRP 27
DB 106 SLSHYGYGQITNEMQKKKAIKFWVSPQWFTAGINP 141

RESULT 13
Q93LZ7 PRELIMINARY; PRT; 420 AA.
AC Q93LZ7;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Chain length factor-like protein.
GN AUR28.
OS Streptomyces aureofaciens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=1894;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CCM3239;
RA Kormanec J., Bistakova J., Novakova R., Homerova D., Rezuchova B.;
RT "Cloning and characterization of a new polyketide gene cluster in
RT Streptomyces aureofaciens CCM3239.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY033994; AAK61719.1; -
DR HSSP; P19435; 1B3N.
DR InterPro; IPR000794; Ketoacyl-synt.
DR Pfam; PF0109; ketoacyl-synt; 1.
DR Pfam; PF02801; ketoacyl-synt; 1.
DR Pfam; PF02801; ketoacyl-synt; 1.
SQ SEQUENCE 420 AA; 43011 MW; 3C27E22BE88C2DEA CRC64;

Query Match 30.2%; Score 46.5; DB 2; Length 420;
Best Local Similarity 52.6%; Pred. No. 41;
Matches 10; Conservative 1; Mismatches 3; Indels 5; Gaps 1;

QY 18 AWYXX-----RGIRPVGRF 31
DB 37 AWAANLARGSGIRPVGRF 55

```

Search completed: December 3, 2003, 19:14:02  
Job time : 26.5 secs

## RESULT 14

Q9UJF9 PRELIMINARY; PRT; 54 AA.  
AC Q9UJF9;  
DT 01-MAY-2000 (TReMBLrel. 13, Created)  
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)  
DE DJ479J7.3 (Sushi-repeat protein (SRPUL)) (Fragment).  
GN DJ479J7.3.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP .SEQUENCE FROM N.A.  
RA Lawlor S.;  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL035608; CAB55682.1; -.  
FT NON\_TER 54 54  
SQ SEQUENCE 54 AA; 6110 MW; E2F3C39F7B961A9F CRC64;

Query Match 29.9%; Score 46; DB 4; Length 54;

Best Local Similarity 43.8%; Pred. No. 5.3;

Matches 7; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 12 TPDINPAWYXXRGIRP 27

DB 18 TPAVPTWVAGSGYYP 33

## RESULT 15

Q8UKM3 PRELIMINARY; PRT; 118 AA.  
AC Q8UKM3;  
DT 01-JUN-2002 (TReMBLrel. 21, Created)  
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
DE Hypothetical protein Atu5077.  
GN ATU5077.  
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).  
OC Plasmid A1.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Rhizobiaceae; Rhizobium.  
OX NCBI\_TaxID=176299;  
RN [1]  
RP .SEQUENCE FROM N.A.  
RX MEDLINE=21608550; PubMed=11743193;  
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,  
RA Okura V.K., Zhou Y., Chen L., Wood G.B., Almeida N.F. Jr., Woo L.,  
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,  
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,  
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A., Gordon D.,  
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Perry M.,  
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,  
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,  
RA Chumley P., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,  
RA Nester E.W.;  
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens C58.";  
RL Science 294:2317-2323(2001).  
DR EMBL; AF008932; AAL45769.1; -.  
KW Hypothetical protein; Plasmid; Complete proteome.  
SQ SEQUENCE 118 AA; 13099 MW; BE71C7BFF8154CF5 CRC64;

Query Match 29.9%; Score 46; DB 16; Length 118;

Best Local Similarity 34.4%; Pred. No. 12;

Matches 11; Conservative 4; Mismatches 13; Indels 4; Gaps 1;

QY 2 RXKHSMEXRTDPI----NPANWYXXRGIRPVG 29

DB 19 KLHKHNAARRRPDYGARVQPATAPHRGLAGIG 50